Revenue Management under Model Uncertainty: Theory and Methods

Omar Besbes

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

COLUMBIA UNIVERSITY

2008
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
A growing number of industries are adopting advanced decision support tools to optimize their revenues. One of the main challenges in the area of revenue management is the ability to account for the underlying uncertainty associated with the demand, e.g., the sensitivity of customers to changes in prices. Most of the literature focuses on cases where the demand model is known and the only uncertainty considered is that associated with random realizations of the demand itself.

This dissertation focuses on revenue management settings in the presence of model uncertainty and in doing so extends the existing revenue management literature along theoretical and practical dimensions. Chapter 2 provides a diagnostic tool to evaluate demand models that are commonly used through a performance based approach that departs from classical statistical approaches. Chapter 3 and 4 focus on single and multi-product dynamic pricing problems under demand model uncertainty. For such problems, we are able to quantify the value of prior information on the demand function as well as to provide near-optimal prescriptions for various levels of uncertainty.
Contents

1 Introduction 1

2 Testing Validity of a Demand Model: an Operations Perspective 13

   2.1 Introduction .......................................................... 13

       2.1.1 Motivation and overview of the main contributions .... 13

       2.1.2 Literature review .............................................. 20

   2.2 Problem Formulation .................................................. 22

   2.3 Background and Model-Based Testing .............................. 24

       2.3.1 Traditional model-based approach and hypothesis testing . 24

       2.3.2 Example of a model-based test ............................... 26

   2.4 The Performance-Based Approach: Key Ideas .................... 28

   2.5 The Performance-Based Approach: Proposed Test ............... 32

       2.5.1 The nonparametric approach ................................. 32

       2.5.2 The test statistic ............................................. 33

   2.6 Properties of the Proposed Test: Illustrative Numerical Examples . 39

   2.7 Empirical Example .................................................... 42

       2.7.1 Data description ............................................. 43
3 The Dynamic Pricing Problem under Model Uncertainty 49

3.1 Introduction ............................................. 49
3.2 Related Literature ......................................... 54
3.3 Problem Formulation ....................................... 57
3.4 Main Results: The Nonparametric Case ..................... 63
  3.4.1 A nonparametric pricing algorithm ....................... 63
  3.4.2 Model uncertainty: the class of demand functions .... 66
  3.4.3 Performance analysis ..................................... 67
3.5 Main Results: The Parametric Case .......................... 70
  3.5.1 The proposed method ..................................... 71
  3.5.2 The parametric class of demand functions .............. 72
  3.5.3 Performance analysis ..................................... 74
  3.5.4 An optimal algorithm when a single parameter is unknown 76
3.6 Numerical Results and Qualitative Insights ................ 79
  3.6.1 The “price” of uncertainty .............................. 79
  3.6.2 The “price” of hedging misspecification risk .......... 81
3.7 Sequential Learning Strategies in the Nonparametric Setting 84

4 Extensions to the Multi-product Case 91

4.1 Introduction ............................................. 91
4.1.1 Background and overview of the main contributions 91
4.1.2 Related literature 96
4.2 Problem Formulation 98
4.3 The Price Restricted Case 103
  4.3.1 The proposed pricing policy 103
  4.3.2 Theoretical analysis 105
  4.3.3 A numerical illustration 107
4.4 The General Case 109
  4.4.1 The blind pricing policy 109
  4.4.2 Performance analysis 113
  4.4.3 A simple state-dependent refinement 115
  4.4.4 Illustrative numerical examples 117
4.5 The Curse of Dimensionality 119
  4.5.1 Exploiting smoothness of the demand function 119
  4.5.2 Exploiting separability of the demand function 127
4.6 Concluding Remarks 128

A Proofs for Chapter 2 139
  A.1 Proof of the Main Result 139

B Proofs for Chapter 3 145
  B.1 Proofs of Main Results 145
    B.1.1 Proofs of the results in Section 3.4 147
B.1.2 Proofs of the results in Section 3.5 153

B.2 Proofs of Auxiliary Results 161

C Proofs for Chapter 4 175

C.1 Proofs of Main Results 175

C.2 Proofs of Auxiliary Results 193
List of Figures

2.1 Demand model misspecification and performance implications . . . 15
2.2 Model misspecification / not rejected by the performance-based test 43
2.3 Model misspecification / not rejected by the performance-based test 44
2.4 Model misspecification / not rejected by the performance-based test 48
3.1 Performance of pricing policies as a function of the market size (n) . 80
3.2 Performance of pricing policies as a function of the market size (n) . 83
4.1 Reconstructing the revenue surface based on price tests . . . . . . 127
List of Tables

2.1 Efficacy of the performance-based test .......................... 40
2.2 Comparison of the performance-based test and model-based test .. 41
2.3 Empirical example: Model-based test vs. Performance-based test . 47
3.1 Performance of stochastic approximation based schemes ............ 89
4.1 Performance in the price restricted case (network) ............... 108
4.2 Performance in the general case (network) ........................ 118
4.3 Exploiting smoothness for performance improvement (network) . . . 125
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Professor Assaf Zeevi for his guidance, encouragements and support at every stage of my doctoral studies. I am also indebted to Professors Costis Maglaras and Sergei Savin for their insights, time and help, and to Robert Phillips at Nomis Solutions for sharing insights and data on various problems. These ultimately led to the second chapter of this dissertation. I thank Professors Guillermo Gallego and Garrett van Ryzin for being part of the thesis committee.

I feel very fortunate to have had the opportunity to interact with all the faculty members and students in the Decision, Risk & Operations division as well as the department of Industrial Engineering and Operations Research. My years at Columbia have definitely been some of the most enriching ones from both academic and personal perspectives.

Thanks also to Elizabeth Elam Chang and Dan Spacher in the doctoral office as well as Barbara Platzer in the Decision, Risk & Operations division for their constant help throughout the process.
Chapter 1

Introduction

Revenue management is a growing subfield of operations research which deals with modeling and optimizing complex pricing and demand management decisions. Since the deregulation of the airline industry in the 1970’s, revenue management practices have become increasingly prevalent in a variety of industries. The literature review in each chapter will also provide a detailed survey of work directly related to this dissertation.

A classical problem in revenue management is the so-called tactical pricing problem: given an initial inventory of products to be sold over a finite selling season, devise a strategy that dynamically adjusts prices so as to maximize the expected total revenues. In this problem it is implicitly assumed that there is little or no control over inventory throughout the time period over which sales are allowed, and pricing is the main lever used to optimize profits. The recent books by Talluri and van Ryzin (2005) and Phillips (2005) as well as survey papers by Elmaghraby and Keskinocak (2003) and Bitran and Caldentey (2003) describe numerous instances of this problem, ranging from fashion and retail, to air travel,
hospitality and leisure. A more complex version of this problem arises when there are several different product types and a set of "resources" (raw materials or primitive components) used to "assemble" them; such problems are often referred to as network revenue management.

A key issue in any revenue management study pertains to the modeling of customer behavior and their reaction to prices. A model vastly adopted in the literature is that of semi-myopic customers. These are assumed to arrive at random points of time with a private willingness-to-pay which is a draw from a given distribution. Upon arrival, a customer purchases the product if her/his willingness-to-pay exceeds the price of the product. In other words, customers are not strategic in the timing of their purchase. Many applications fall under this category and only such settings will be considered in this dissertation. It is worth noting that there are instances where customers might strategically decide on the timing of their purchase and that this leads to fundamentally different considerations. Such cases are not discussed in this dissertation and we refer the reader to, e.g., Shen and Su (2007) for an overview of that line of research.

In the context we focus on, a crucial concept is the functional relationship between the mean demand rate and price, often referred to as the demand function or demand curve. One can think of this as an aggregate characterization of the market that is derived from the willingness-to-pay distribution which governs customer behavior; see Talluri and van Ryzin (2005), Phillips (2005) and Train (2002) for further discussion on customer behavior modeling. A critical assumption made
in most academic studies of revenue management problems is that the demand function is known to the decision maker. As a result, the only form of uncertainty is due to randomness of demand realizations. This makes the underlying problems more tractable and allows one to extract structural insights. At the same time, this assumption of "full information" endows the decision maker with knowledge that s/he does not typically possess in practice. It is far more realistic to assume that a firm possesses only limited information with regard to the demand function. This issue is especially relevant for fashion items and technical products with short life cycles, but it also arises in many other instances in which limited historical sales data does not allow for accurate inference of the price-demand relationship. In such cases one can view the assumption of full information as more of a convenient mathematical abstraction (which facilitates studying the structural properties of optimal pricing policies), rather than an accurate description of the information available to the decision maker.

The presence of uncertainty associated with the demand function gives rise to several important questions, both from a fundamental theoretical perspective, as well as from an applications standpoint. There are various levels of uncertainty that firms might be facing. In the presence of historical data, firms can already make inferences based on past observed purchasing behavior. In these instances, it is natural to ask how should one use available data to guide decision-making or how good are "classical" demand models that firms typically rely on in practice. In the absence of historical sales data, the problem faced faced by the firm becomes
more challenging as there is no basis from which to infer information initially. As a result, information gathering becomes part of the overall revenue optimization problem. The design of prescriptions in such settings, in conjunction with tools to assess their quality, are important research questions.

This dissertation seeks to address some of the above questions. In particular, we highlight the general interplay between estimation, learning and optimization in revenue management. The dissertation is organized around three main parts.

**Chapter 2.** Chapter 2 analyzes a pricing problem where the demand function is unknown but the firm has access to historical sales data. The main question addressed there pertains to the validity of a given class of demand models in the context of revenue optimization. While the statistics and econometrics literature have developed powerful methods for testing the validity (specification) of a model, managers are typically more interested in the performance of their decisions rather than the validity of the model from which they are derived. Focusing on the problem of demand model specification in the context of revenue management, we propose a framework and a statistical test that captures this perspective. Theoretical properties of the test are established and its efficacy is illustrated both on synthetic examples, as well as on an empirical data set in the realm of financial services. It is shown that traditional model-based goodness-of-fit tests may consistently reject simple parametric models of consumer response (e.g., the logit model), while at the same time these models may "pass" the proposed performance-based test.
An important takeaway of the second chapter is the importance of considering the ultimate problem of interest (here, the downstream optimization) when analyzing model specification. It will be apparent that the ideas developed in the context of revenue management are pertinent to other operations management problems. To the best of our knowledge, the proposed test is the first one that strives to assess the validity of a model from purely operational considerations, i.e., focusing on decision making implications of model specification, and shifting the focus away from a statistical perspective. The only other study we are aware of that shares some common ground with this theme is Liyanage and Shantikumar (2005), who examine demand model uncertainty effects on the design of inventory policies. They do not focus on model testing or specification per se, nor on statistical analysis, and therefore the intersection of their paper and this work is more in terms of philosophy rather than the main focus and methods that are developed.

The test presented in Chapter 2 allows for a direct implementation, and can hence provide a useful diagnostic tool for practitioners. Interestingly, while the approach we develop relies on nonparametric estimation methods, the test itself may actually be used to lend support to parametric modeling. In particular, it can help identify whether simple parametric families support near-optimal decision making despite being potentially misspecified. Given that parametric families will often be rejected using traditional model-based tests, their justification from an operations perspective may be one of the more useful outcomes of our proposed test.
The premise of the analysis in Chapter 2 is the availability of historical sales data. In the absence of the latter, a firm is facing a starker challenge as there is no clear basis point that one can use to devise a potential demand model. The rest of the dissertation focuses on the more general problem of dynamic pricing problems without knowledge of the demand function and in the absence of historical data. We anchor our analysis around the prototypical revenue management problem discussed earlier: given an initial level of inventory and a finite selling horizon, dynamically price the products being sold so as to maximize the total expected revenues. The implication of demand model uncertainty is that this class of problems is characterized by an intrinsic tension between exploration (demand learning) and exploitation (pricing/optimization).

**Chapter 3.** The problem just described was first formalized in Gallego and van Ryzin (1994) as an intensity control problem in the context of “full information,” i.e., when the demand function is known. Chapter 3 considers the problem above in cases where the demand function is *not known*. We focus on two basic levels of uncertainty with regard to the demand model: i.) the *nonparametric setting* where the demand function is only assumed to belong to a broad functional class satisfying mild regularity conditions; and ii.) a *parametric setting* in which the demand function admits a given parametric structure but the parameter values are not known to the decision maker. Our main goal is to introduce suitable algorithms that learn the demand function “on the fly,” and use that as a basis for pricing decisions. Their performance will be measured in terms of the revenue loss.
relative to a full information benchmark that assumes knowledge of the demand function. We refer to this loss as the *regret* associated with not knowing the demand function a priori; the magnitude of the regret quantifies the *economic value* of prior model information. The policies we consider are designed with the objective of achieving a "small" regret *uniformly* over the relevant class of demand functions (either parametric or nonparametric). This adversarial setting, where nature is allowed to counter a chosen policy with the "worst" demand function, ensures that policies exhibit "good" performance irrespective of the true demand model.

The complexity of the problem described above makes it difficult to evaluate the performance of any reasonable policy, except via numerical experiments. To address this issue, we consider an asymptotic regime which is characterized by a high volume of sales. More specifically, the initial level of inventory and the magnitude of demand ("market size") grow large in proportion to each other. This regime allows us to bound the magnitude of the regret, and to characterize the performance of our proposed policies.

Relatively few studies in the revenue management literature consider uncertainty with regard to the demand function, and most of this line of work is pursued within the context of a parametric structure in which one or more parameters are not known. The typical approach there is to use a dynamic programming formulation with Bayesian updating, where a prior on the distribution of the unknown parameters is initially postulated. Recent examples include Aviv and Pazgal (2005),
Araman and Caldentey (2005) and Farias and Van Roy (2006), who all assume the market size parameter is unknown. (See also Lobo and Boyd (2003) and Carvalho and Puterman (2005).) While the Bayesian approach is attractive as it allows for a stylized analysis of the joint learning and pricing problem, it suffers from significant shortcomings. Most notably, the objective of the dynamic optimization problem involves an expectation that is taken relative to a prior distribution. Hence any notion of optimality associated with a Bayesian-based policy is with respect to that prior.

The analysis of parametric settings in Chapter 3 contrasts with the above papers as we adopt a "frequentist" approach based on Maximum Likelihood estimation. We propose pricing algorithms and establish that as the sales volume grows large, the regret eventually shrinks to zero. That is, these policies achieve (asymptotically) the maximal full information revenues, despite the absence of prior information regarding the demand function; in that sense, they are asymptotically optimal. In addition, we derive asymptotic lower bounds on the regret that hold for any admissible pricing policy. In light of these bounds, the performance of the proposed algorithms is "not far" from optimal in the sense of the magnitude of the regret, and in some cases cannot be improved upon (i.e., no admissible policy can achieve a smaller regret).

While the proposed parametric approach circumvents some of the problematic points in the Bayesian formulation, it still shares a very significant shortcoming with the latter, as well as with any parametric approach. In particular, for the
method to work well, it is crucial that the structure assumed by the policy be consistent with that of the true underlying demand function. In other words, the postulated model needs to be well specified with respect to the actual mechanism generating demand observations. In the absence of this premise, any parametric policy may exhibit extremely poor performance. To counter such misspecification risks, we investigate settings where the decision-maker does make any a priori assumption about the form of demand function; this is the nonparametric setting. In the latter, we propose a class of pricing policies that separates estimation (exploration / demand learning) and control (exploitation / pricing to maximize revenues). Their broad structure runs as follows. In the first phase (exploration), one uses "test prices" to gather data on the underlying model. This information is then used to construct an empirical version of a deterministic relaxation of the original stochastic finite horizon revenue maximization problem. Finally, the solution of this optimization problem gives rise to the ultimate pricing policy that will be used immediately following the learning phase, throughout the remainder of the selling season (exploitation). We show that by judiciously selecting the number of prices used to experiment with as well as the length of the experiments, the proposed policies are asymptotically optimal. In other words, the value of knowing the demand function a priori shrinks as the sales volume grows large.

Overall, Chapter 3 develops lower and upper bounds on the magnitude of the regret for different levels of model uncertainty. These can be used to rigorously quantify the economic value of a priori information on the demand model, by
assessing the revenue loss that results from incomplete information; the "price
of model uncertainty." In addition, they provide a means for quantifying the
"price" that one pays for eliminating the misspecification risk (discussed earlier)
via nonparametric approaches.

Chapter 4. Chapter 4 is aimed at extending the previous analysis to network
revenue management settings under demand model uncertainty. In particular, we
focus exclusively on nonparametric settings in this chapter.

The full information version of the problem was formulated and analyzed in
Gallego and van Ryzin (1997). In the presence of demand model uncertainty, the
dimensionality of the network problem presents key challenges in the execution of
the ideas developed in the context of the single product problem in Chapter 3.
Recalling the structure of the proposed policies in the nonparametric setting, two
main main issues would appear here. First, on the exploration phase, one needs to
select a suitable and sufficient rich set of "test prices" for purposes of learning the
multidimensional nonparametric demand function. Second, for purposes of the
exploitation phase, one needs to properly formulate a constrained optimization
problem to derive the ultimate pricing strategy.

We first develop a naive policy that tests a discrete set of prices in the ex­
ploration phase, and then selects the "best" price to be used in the exploitation
phase. Our analysis establishes that the policy is asymptotically optimal, but its
performance degrades significantly with the number of products being sold. This is
a manifestation of the curse of dimensionality. With this observation in mind, we
then propose a modification of this policy that uses the demand data obtained at
the price testing phase to construct an estimate of the entire demand surface. This
functional estimate is then fed into a proper deterministic optimization problem
which gives rise to the ultimate pricing policy. The key idea here is to use prior
knowledge on the smoothness of the demand function to guide the data collection,
and nonparametric curve fitting stages. Roughly speaking, the smoother the de­
mand surface, the less one suffers from dimensionality effects; this is articulated in
precise mathematical terms in the analysis. When the unknown demand function
is suitably smooth, we exhibit blind pricing policies that build on ideas in non‐
parametric statistics whose performance degrades gracefully with the dimension
of the problem, effectively mitigating the curse of dimensionality. An important
implication of this is that good performance can be achieved while requiring only a
moderate number of prices to be tested in the learning phase, making the approach
more appealing from a practical perspective.

A useful insight that arises from the analysis of Chapters 3 and 4 is related
to the industry practice of “price testing,” a prevalent method used by firms to
address the lack of precise demand information; a recent empirical study of 32 large
U.S. retailers, finds that nearly 90% of them conduct price experiments (see Gaur
and Fisher (2005)). The main idea is quite straightforward and closely related
in spirit to the structure of our algorithm: in the first step one experiments with
several prices; and in the second step one selects the price/s that are expected to
optimize revenues based on the demand observed in the previous step. Among
the main questions that arise in this context is how many prices to test, and for how long. Current practices are mostly guided by ad hoc considerations in addressing such issues. Given the significant role that price testing plays in revenue management practices, there is a growing need to improve the understanding of this approach and add to its rigorous foundations (see, e.g., Williams and Partani (2006) for further discussion and examples). Our analysis contributes to this goal by providing simple and intuitive guidelines for selecting both the number of prices that should be tested, as well as the overall fraction of the selling season that should be dedicated to experimentation.

Summary of main contributions. The dissertation extends the existing revenue management literature along the dimension of decision-making under model uncertainty. Chapter 2 provides a diagnostic tool that uses historical sales data to evaluate demand models that are commonly used by adopting an operations perspective as opposed to a purely statistical one. Chapters 3 and 4 analyze at single and multi-product dynamic pricing problems under demand model uncertainty. For such problems, we are able to quantify the value of prior information on the demand function as well as to provide near-optimal prescriptions for various levels of uncertainty.
Chapter 2

Testing the Validity of a Demand Model: an Operations Perspective

2.1 Introduction

2.1.1 Motivation and overview of the main contributions

Understanding the structure of realized demand and the underlying consumer purchasing behavior plays a key role in almost all areas of revenue management. In the academic literature on pricing and capacity allocation problems, demand is typically modeled in a manner that facilitates stylized analysis from which qualitative insights are derived. In contrast, in most real world settings prescriptive solutions are typically sought. In the presence of historical sales data, a possible (and prevalent) approach follows along these lines: (i) specify a parametric family to be used as the demand model; (ii) estimate its parameters using available sales data; and (iii) optimize the calibrated model with respect to decision variables. (See Talluri and van Ryzin (2005) and Phillips (2005) for examples and further pointers to the literature.)

The successful application of the approach outlined above would require that
the postulated family of demand models matches, in some sense, the true underlying mechanism that determines realized demand. Thus a natural step that should be inserted prior to the optimization stage (iii) above is to assess the validity (or equivalently, misspecification) of the postulated family of models. These type of questions have been studied extensively in the fields of statistics and economics (econometrics). The focus there has been, by and large, on methods for testing the hypothesis that a proposed class of models (typically a parametric family) contains the true structure that generates the data. In other words, the tests examine whether the specification of the model class is correct; specific references pertinent to the present chapter will be given in Section 2.1.2.

To motivate the approach we develop in this chapter, and illustrate how it differs from the traditional objective outlined above, consider the following simple set up which is characteristic of many pricing problems encountered in practice. Consumers' purchase decisions are governed by a so-called response function or equivalently, a willingness-to-pay distribution, denoted \( \lambda(x) = \text{probability of purchasing the product at a price } x \). (This function is assumed not to change over the relevant time horizon over which decision are made.) The revenue manager's objective is to maximize the expected profit-per-customer, say, \( \pi(x) = (x - x_0)\lambda(x) \), by suitably setting the price variable \( x \). The main issue is that the true underlying response function is not known, and for this reason a class of (demand) models is put in place, based on which the profit maximization objective is solved. The question then is:
What is the loss incurred by using pricing decisions derived from this restricted class of models relative to the best achievable performance?

If the answer is that this loss is suitably “small,” the assumed demand model, whether well specified or not, might be deemed adequate insofar as it leads to pricing decisions that generate near-optimal profits.

Figure 2.1: **Demand model misspecification and performance implications.** $\Delta =$ optimal profit rate minus profits achieved using model based price-estimate. The magnitude of $\Delta$ quantifies the loss due to model misspecification.

In Figure 2.1(a), we depict two demand models (response functions) that characterize substantially different consumer choice characteristics. In Figure 2.1(b), the profit function associated with each model is represented. Suppose that the true response function characterizing consumer demand is given by the full line, and that the postulated model used to derive the pricing decision is depicted by
the dotted line. The optimal price, relative to the true response function [solid line] would be $5. The "best" price prescribed by the assumed demand model [dotted line] is $6.

Figure 2.1(b) depicts the difference $\Delta$ between the optimal profit rate, and the profit rate realized by the model-prescribed decision ($x = 6$); the latter describes the profits that are achieved using this prescription, that is, when consumer behavior is dictated by the true underlying response function. The main idea is that if $\Delta$ is suitably small, so that based on a sample of past sales it cannot be distinguished from zero in a well defined statistical sense, then the decision maker in our problem will be content with the performance of the misspecified model s/he is using. This view represents a significant departure from the traditional statistical perspective on such problems: the two demand models in Figure 2.1(a) are easily distinct, and most reasonable statistical model-based tests would end up rejecting the notion that they are identical based on sufficient data generated by any one of them. In contrast, an operations perspective should focus directly on the discrepancy in performance.

The main contribution of this chapter is to place the observations made above on rigorous grounds. This is done by developing a hypothesis test that allows one to assess the validity of a model based on its performance implications. To be concrete, we focus on a revenue management setting outlined in broad strokes above, but it should be apparent that the ideas are pertinent to other operations management problems. The performance-based test we propose is effectively aimed
at quantifying what values of $\Delta$ (the loss in profits due to the use of a given model class) differ from zero in a statistically significant manner. To meet that goal, we prove that the distribution of an appropriate estimator of $\Delta$, suitably normalized, converges to a simple limit distribution as the sample size grows large; see Theorems 2 and 3. The limit is a scaled Chi-squared distribution and several numerical examples are developed to illustrate the behavior of the test statistic and the consistency of the test. We also illustrate the applicability of these methods via an empirical study that originates in a financial services application (see further discussion below).

The limit theory alluded to above allows us to quantify what values of performance loss can be considered significant, in a precise statistical sense, so as to reject the family of postulated demand models based on observed data. To the best of our knowledge, this is the first such test that strives to assess the validity of a model from purely operational considerations, i.e., focusing on decision making implications of model specification.

In terms of methodology, we rely on nonparametric estimation techniques to establish a consistent estimator of the true unknown response function, and relevant asymptotic theory for such estimators. Large sample theory of maximum likelihood estimation in a misspecified environment also plays an important role in our derivations; see, e.g., White (1996) and the literature review below for general references on these topics.

**Origin of the research questions and bearings on practice.** This re-
search was initiated when analyzing pricing data for a financial services firm that offers automobile loans. Each customer was offered a rate for a given term of loan, and the ultimate response (accept/reject) was recorded; further details on the nature of the data are discussed in Section 2.7. The initial objective was to assess whether the response function that was used to model the consumer's decisions was adequate. In particular, the main question was whether the firm should continue using the hypothesized parametric model (which consisted of a logit response function with estimated parameters), or fundamentally change this model.

An initial analysis of the data revealed that certain values of the decision variable (offered rate) that were tested led to very poor performance. In some sense, one would really like to focus on the "price sensitivity" only in the neighborhood of the point of maximum of the profit function. A systematic way of ignoring "less relevant" regions in the decision variable space would be the key to operationalizing such a test. This led to the performance-based approach that is proposed in this chapter. Our proposed test allows revenue managers to assess if historical demand data supports the use of a given model focusing on the performance of decisions that it breeds. Roughly speaking, rather than globally testing the entire demand model, it effectively localizes the test to a region of interest (at which near-maximal revenues are obtained). The test presented in the current manuscript together with the supporting limit theory allows for a fairly straightforward implementation, and can hence provide a useful diagnostic tool for practitioners.

Interestingly, while the approach we develop relies on nonparametric estimation
methods, the proposed test may actually be used to lend support to parametric modeling. In particular, it can help identify whether simple parametric families support near-optimal decision making despite being potentially misspecified. For practitioners this, most likely, would be a desirable outcome given the complexity and non-transparent nature of nonparametric methods (fact to point, these methods are very rarely used in industry). Parametric models, in contrast, admit a straightforward interpretation, and can be easily used to assess price sensitivity, elasticity and other important demand-related characteristics. Given that parametric families will often be rejected using traditional model-based tests (see Sections 2.6 and 2.7), their justification from an operations perspective may be one of the more useful outcomes of our proposed test.

The remainder of the chapter. In the next section we provide a review of literature that is most closely related to our chapter, and pointers for further background reading on response function modeling, specification testing, nonparametric estimation, and other topics that intersect in our work. In Section 2.2 we formulate the revenue management problem. Section 2.3 provides some background on model-based testing and hypothesis testing. Section 2.4 introduces the performance-based approach; Section 2.5 presents a new test and its properties which are then numerically illustrated in Section 2.6. Section 2.7 presents an application of the test to an empirical example. Appendix A.1 provides the proofs of the main results.
2.1.2 Literature review

Our focus in this chapter is on a revenue management application which centers on consumer choice behavior. There is a large stream of literature focusing on such models. Ben-Akiva and Lerman (1985) and Train (2002) are examples of such that provide an overview on the topic. In a large number of applications, firms would fit a number of parametric models and compare them based on some criteria before deciding on the "best" one. This is based on a significant body of work, dating back to the pioneering paper of Akaike (1974), that deals with model selection; the typical approach there is to formulate an optimization problem that penalizes the complexity of a model (e.g., log-likelihood with penalty for the number of parameters in the model). It is important to note that all models considered in these comparisons may still be misspecified with respect to the true mechanism that generates the data. Amemiya (1981, Section II.C) reviews some criteria typically used for model selection, and gives a general econometrics perspective on this and related issues; Leeftang et al. (2000) provide a general overview from a marketing perspective.

Distinct from that line of research is the model testing paradigm in which the specification of a model is tested against the true underlying structure of interest; this approach is covered in almost any graduate level textbook on statistical theory; see, e.g., Borovkov (1998). Broadly speaking, our work falls into the latter category. For the purpose of our revenue management application, the response function is a conditional probability and hence one can draw on general results that
have been developed in the literature for testing the validity of regression (conditional expectation) type models. Perhaps the two most notable examples are the conditional moment tests of Bierens (1990), and the conditional Kolmogorov test of Andrews (1997); see also references therein for further pointers to this literature. Both these tests enjoy certain optimality properties in terms of their power against local alternatives. Unlike these types of tests, that do not directly estimate the regression function, there are various others that use intermediary nonparametric approximations to the regression function; see, e.g., Härdle and Mammen (1994).

Various instances of the model specification tests detailed above have been applied to empirical data sets in order to assess the validity of widely used models such as the Logit or Probit. Horowitz (1993) analyzes the binary response model of choice between automobile and public transit, and Bartels et al. (1999) apply a nonparametric test to scanner panel data and reject the multinomial logit model; see also references therein. While not focusing on testing per se, Abe (1995) provides a discussion that focuses on the benefits and drawbacks of nonparametric models relative to simple parametric models (logit) in the context of marketing research.

The current paper provides a novel goodness-of-fit test, and applies it to an empirical dataset. We rely heavily on nonparametric estimation methods, accessible overviews of which are Härdle (1990) and Pagan and Ullah (1999). The distinguishing feature of our study is that it shifts the focus from a statistical to an operational perspective. The only other study we are aware of that shares some
common ground with this theme is Liyanage and Shantikumar (2005), who examine demand model uncertainty effects on the design of inventory policies. They do not focus on model testing or specification per se, nor on statistical analysis, and therefore the intersection of their paper and ours is more in terms of philosophy rather than the main focus and methods that are developed.

2.2 Problem Formulation

We consider a problem where a single product can be sold for a price $x \in \mathcal{X} := [\underline{x}, \bar{x}]$, with $0 < \underline{x} < \bar{x} < \infty$. At the prevailing price, $x$, a consumer will purchase the product with probability $\mathbb{P}(Y = 1|x)$; where $Y \in \{0, 1\}$ is a random variable such that $Y = 1$ corresponds to a purchase decision; and $Y = 0$ corresponds to a situation where the consumer declines to purchase the product. We refer to $\lambda(x) := \mathbb{P}(Y = 1|x)$ as the consumer response function.

Let $r(x)$ denote a function that describes the revenue/profit resulting from a given sale. The decision-maker's objective is to set a price that maximizes the expected profit per customer. That is, for the profit function

$$\pi(x) := r(x)\lambda(x),$$

the objective is to seek $x^* \in \text{arg max}\{\pi(x) : x \in \mathcal{X}\}$. Under mild conditions on $\pi$, e.g., continuity, such a point of maximum exists. The optimal profits are then $\pi^* := \pi(x^*)$.

The decision-maker does not know the true response function $\lambda(\cdot)$ characterizing the market. S/he only has access to data in the form of $n$ past observa-
tions $\mathcal{D}_n = \{(X_i, Y_i) : 1 \leq i \leq n\}$; each pair $(X_i, Y_i)$ describes the sales outcome $Y_i \in \{0, 1\}$ when offering the product at price $X_i \in \mathcal{X}$. Using this data, a model for the response function is fitted from the parametric family $\mathcal{L}(\Theta) = \{\ell(\cdot; \theta) : \theta \in \Theta\}$, where $\Theta \subseteq \mathbb{R}^d$ is a compact set. The notation here is indicative of the fact that the postulated model for the response function belongs to a parametric family, which is the most widely used approach in practice (as discussed in the introduction). In what follows, we use $\mathbb{P}_\theta(y|x)$ to denote the conditional probability of observing $y$ when $x$ is chosen under the model $\ell(\cdot; \theta)$. For any $\theta \in \Theta$, we let $p(x; \theta) := r(x)\ell(x; \theta)$ denote the profit function under the parametric assumptions describing the postulated model. Put $x^*(\theta)$ to be a maximizer of $p(x; \theta), x \in \mathcal{X}$. Hence for a fixed choice of the parameter $\theta \in \Theta$, $x^*(\theta)$ represents the optimal model-based decision price. In practice, the value of the parameter $\theta$ is estimated from the data $\mathcal{D}_n$ and let $\hat{\theta}$ denote such an estimate, i.e., a mapping from $\mathcal{D}_n$ to $\Theta$. While many such estimation procedures exist, depending on the context, we will focus here on that of maximum likelihood estimation, which is by far the one most prevalently used in practical settings.

**The main question.** The method outlined above arrives at a prescribed price as follows: estimate $\hat{\theta}$, compute $x^*(\hat{\theta}) \in \text{arg max}\{p(x; \hat{\theta}) : x \in \mathcal{X}\}$. The classical statistical model-based approach would focus on whether $\ell(\cdot; \hat{\theta})$ is a good approximation to the true response function $\lambda(\cdot)$; roughly speaking, whether $\ell(\cdot; \hat{\theta}) \approx \lambda(\cdot)$. As indicated above, there is a significant statistical literature that develops rigorous methods to test such a hypothesis. In contrast, a revenue manager is interested
in the performance of the prescribed price, namely, how \( \pi(x^*(\theta)) \) relates to \( \pi^* \). In particular, it is natural to consider \( x^*(\theta) \) to be near-optimal if \( \pi(x^*(\theta)) \approx \pi^* \). The focus of this paper is to develop an operations-based perspective that focuses on the performance of the decision prescribed by the model class \( \mathcal{L}(\Theta) \), rather than on the model per se.

### 2.3 Background and Model-Based Testing

We first describe briefly the classical model-based approach which is the one often found in the statistics and econometrics literature. In doing so, we also introduce some necessary background on hypothesis testing and the key concepts that will be used throughout the paper. Subsequent to that, we review briefly a model-based test developed by Andrews (1997), that will later serve as a basis for comparison against our proposed performance test which is described in Section 2.5.

#### 2.3.1 Traditional model-based approach and hypothesis testing

The traditional statistical approach strives to determine, based on observations \( D_n \), whether the true unobservable conditional probability \( \lambda(\cdot) \) can be distinguished from the "best approximation" within the model class \( \mathcal{L}(\Theta) \). Formally, one can formulate the hypothesis test as follows:

\[
H_0 : \quad \lambda(\cdot) = \ell(\cdot; \theta_0) \quad \text{for some } \theta_0 \in \Theta \\
H_1 : \quad \lambda(\cdot) \neq \ell(\cdot; \theta) \quad \text{for all } \theta \in \Theta,
\]
where the ‘≠’ in the alternative hypothesis means that for any $\theta \in \Theta$ there exists some $x \in [x, \overline{x}]$ such that $\lambda(x) \neq \ell(x; \theta)$. It is worth emphasizing that what one is testing via this traditional statistical formulation is a hypothesis about the model that generates the data.

The decision rule that is used to resolve the test typically hinges on a suitably chosen test statistic $T_n : D_n \to \mathbb{R}_+$. The idea is that when $T_n$ (properly scaled) exceeds, say, a suitably chosen threshold $\tau$, the null hypothesis $H_0$ is rejected. Since the distribution of “good” test statistics is often difficult to compute, one resorts to an asymptotic analysis. (Note also that the hypotheses in (2.3.1)-(2.3.2) are not simple hypotheses, in the sense that the distribution of the response $Y$ conditional on the covariate $X$ is not fully specified under $H_0$.)

Define a scaling sequence of positive real numbers $\{a_n\}$ such that $a_n T_n$ converges in distribution to a limit random variable $Z$; we denote this as $a_n T_n \Rightarrow Z$ as $n \to \infty$. The threshold $\tau$ is then chosen so that

$$P\{a_n T_n > \tau \mid H_0\} \to \alpha \quad \text{as } n \to \infty, \quad (2.3.3)$$

where $\alpha \in (0, 1)$ is called the significance level of the test. That is, the choice of $\tau$ ensures that the Type 1 probability of error, i.e., the likelihood of rejecting the null when it is true, is asymptotically equal to $\alpha$. A decision rule or test (we use the two interchangeably in what follows) is said to be consistent if

$$P\{a_n T_n \leq \tau \mid H_1\} \to 0 \quad \text{as } n \to \infty. \quad (2.3.4)$$

This restriction ensures that the Type 2 error, namely, not rejecting the null when
it is false, has vanishingly small probability as the sample size increases. Finally, the \textit{p-value} associated with the test is the minimum level of significance at which one fails to reject the null hypothesis: if the \textit{p-value} falls below the significance level \( \alpha \), the null is rejected (the smaller the \textit{p-value}, the more evidence there is to reject the null).

Testing model specification as in (2.3.1)-(2.3.2) has received significant attention in the fields of economics (econometrics) and statistics as discussed in Section 2.1. A specific test that falls into this category and has certain desired properties is described in the next section.

\subsection*{2.3.2 Example of a model-based test}

Below, we briefly describe a model-based test developed by Andrews (1997) which resembles the classical Kolmogorov-Smirnov test used to assess if two samples are drawn for the same distribution (see, e.g., Borovkov (1998)). Based on historical data \( D_n = \{ (X_i, Y_i) : 1 \leq i \leq n \} \), the test uses the conditional Kolmogorov test statistic which is defined as follows

\[ CK_n = \sqrt{n} \max_{j \leq n} \left| \frac{1}{n} \sum_{i=1}^{n} 1\{Y_i \leq Y_j\} 1\{X_i \leq X_j\} - \frac{1}{n} \sum_{i=1}^{n} \mathbb{P}_{\theta}(Y_i \leq Y_j | X_i) 1\{X_i \leq X_j\} \right|, \]

where \( \hat{\theta} \) is the maximum likelihood estimate of \( \theta \) and in our set up \( \mathbb{P}_{\theta}(Y_i \leq Y_j | X_i) = 1 \) if \( Y_j = 1 \) and \( \mathbb{P}_{\theta}(Y_i \leq Y_j | X_i) = 1 - \ell(X_i; \theta) \) if \( Y_j = 0 \). The Conditional Kolmogorov test statistic compares two terms. The first is an empirical version of the cumulative distribution function of the vector \((X, Y)\). As the sample size grows large, this term converges to the true cumulative distribution function (cdf) of the
vector \((X, Y)\) by the law of large numbers. The second term is a semi-parametric, semi-empirical version of the cdf of \((X, Y)\), and under the null hypothesis, it would also converge to the same limit as above. However, when the null is false, the two limits will differ, and the statistic, which is scaled by \(\sqrt{n}\) will diverge to infinity.

In that regard, we make the following assumption

**Assumption 1** For \(i \geq 1\), \((X_i, Y_i)\) are independent and identically distributed (iid) with response function given by \(\mathbb{P}(Y = 1|X = x) = \lambda(x)\) and marginal distribution \(G\) with density function \(g(\cdot)\) which is positive everywhere on its support \([x, \bar{x}]\).

Assumption 1 is adopted for convenience as it facilitates the mathematical analysis in what follows (the derivation of large sample properties of certain test statistics is the focal point). From that perspective, one can significantly weaken this assumption and still attain the type of limit theory allowing for mild dependency of \((X_i, Y_i)\) pairs, and this is well documented in the statistical literature (see, for example, White (1996)). From a practical perspective, we require this assumption mostly to ensure sufficient dispersion of the decision variable \((X)\). In the absence of this, it would be impossible to reconstruct the response function consistently.

In the empirical study presented in Section 2.7, we will illustrate that such price dispersion is present in the data (see Figure 2.4). This is typical of cases where price experiments are conducted to infer the nature of the consumer purchasing decision. We also make the following technical assumption
Assumption 2  Each member of the family of response function models \( \ell(x; \theta) \in \mathcal{L}(\Theta) \) is continuously differentiable on \([a, b] \times \Theta\).

We state below the result for the special case under consideration here.

Theorem 1 (Andrews (1997)) Let Assumptions 1 and 2 hold. Under the null hypothesis,

\[ CK_n \Rightarrow \mathcal{K} \quad \text{as } n \to \infty. \]

The precise characterization of the limit \( \mathcal{K} \) is given in Andrews (1997). The limit distribution depends on several nuisance parameters including the true parameter vector \( \theta^* \) and the marginal distribution \( G(\cdot) \). In addition, the form of the limit distribution is fairly complicated. To implement the test, Andrews (1997) suggests a bootstrapping procedure. Further details on the implementation of this test will be discussed in sections 2.6 and 2.7.

2.4  The Performance-Based Approach: Key Ideas

Motivation. The model-based approach, and hence the test described in (2.3.1)-(2.3.2), focuses on the specification of the response model. Roughly speaking, for a given estimator \( \hat{\theta} \), the null hypothesis will be rejected when the estimated response function \( \ell(\cdot; \hat{\theta}) \) differs in a statistically significant manner from the true response model \( \lambda(\cdot) \). Yet even under such circumstances it is possible that the estimated response function would still give rise to a "good" pricing prescription, i.e., a price that performs well under the true underlying response function. As illustrated in
Figure 2.1 in the introduction, this is possible even if the postulated model class $\mathcal{L}(\theta)$ is completely misspecified relative to the underlying response model $\lambda(\cdot)$.

We now describe a new test that focuses directly on the performance of decisions derived from the parametric model class $\mathcal{L}(\theta)$. In particular, the question that will be answered by this test is whether decisions that are derived from the "best" parametric model in $\mathcal{L}(\theta)$ lead to actual performance (profits) that differ significantly from (i.e., are inferior to) the best achievable performance. The latter corresponds to profits generated by the optimal decision derived from the true underlying response model.

**Background on parametric inference and model misspecification.** To describe our proposed test, we first need to articulate what is meant by the "best" parametric model in $\mathcal{L}(\theta)$, as that class need not include the true response function $\lambda(\cdot)$; a good reference on the topic is the book by White (1996).

By far the most widely used method for fitting a parametric model to given data is that of maximum likelihood. Let

$$\hat{\theta} \in \arg \max \left\{ \frac{1}{n} \sum_{i=1}^{n} \log \left( \mathbb{P}_\theta(Y_i|X_i) \right) \right\},$$

be the maximum likelihood estimator based on the sample $\mathcal{D}_n$, where $\mathbb{P}_\theta$ denotes the conditional probability distribution of $Y$ given $X$ for a parameter vector $\theta \in \Theta$. In our context, $\mathbb{P}_\theta(y = 1|x) = 1 - \mathbb{P}_\theta(y = 0|x) = \ell(x; \theta)$. The right-hand-side above, which is being maximized, is a sample-based approximation to the expected log-likelihood $E\left[ \log \left( \mathbb{P}_\theta(Y_i|X_i) \right) \right]$, where the expectation is with the respect to the
true distribution of \(X_i\)'s and \(Y_i\)'s, which may be distinct from any distribution \(\{P_\theta : \theta \in \Theta\}\). It is easily seen that the expected log-likelihood is maximized for a value of \(\theta = \theta^*\) which minimizes

\[
\mathbb{E} \left[ \log \left( \frac{P(Y_i|X_i)}{P_\theta(Y_i|X_i)} \right) \right],
\]

over the parameter space \(\Theta\).

The expression in (2.4.1) is called the Kullback-Leibler Information Criterion (KLIC) and can be shown to be non-negative and equal to zero if and only if \(\theta\) is such that \(\lambda(\cdot) = \ell(\cdot; \theta)\) for almost all \(x\) in its support \(\mathcal{X}\). Hence one can think of KLIC as a measure of "distance" between the true underlying response function, and members of the parametric class \(\mathcal{L}(\Theta)\) (although KLIC is not formally a metric since it does not satisfy the triangle inequality). As \(n\) grows large, the empirical log-likelihood converges to the expected log-likelihood, and hence one expects that \(\hat{\theta}\) will converge to the point \(\theta^*\) which minimizes the KLIC. The latter can be shown to hold under some mild technical conditions (see White (1996)). Thus, when using ML estimation, one is effectively using a finite sample approximation to the parameter that minimizes the KLIC measure of distance of the parametric model class \(\mathcal{L}(\Theta)\) from the true underlying response function which generates the data.

**Formulation of the performance-based test.** The Hypothesis test can be
written as follows

\[ H_0 : \quad \pi^* = \pi(x^*(\theta^*)) \quad (2.4.2) \]

\[ H_1 : \quad \pi^* > \pi(x^*(\theta^*)) , \quad (2.4.3) \]

where: \( \pi^* \) represent the optimal profit rate; \( x^* \in \arg\max\{\pi(x) : x \in \mathcal{X}\} \) is the optimal price relative to the true underlying profit function; and \( x^*(\theta) \) is the maximizer of the profit rate corresponding to the parametric model class \( p(x; \theta) = r(x)\ell(x; \theta) \) for \( \theta \in \Theta \). Thus what is being compared above is the best achievable performance \( \pi^* \), and the performance achieved by a decision (price) that optimizes the "best fit" parametric model.

An important observation here is that the new test is quite different in flavor from the traditional statistical one given in (2.3.1)-(2.3.2). The latter would end up rejecting a given parametric model class unless it provides a "good" global fit to the true underlying response function. The test described above is local in nature: it will reject the null only if the resulting price prescriptions do not fall within the region where the true profit function achieves its maximum. In this manner, we substantially relax the definition of the null \( H_0 \) in comparison with the model-based test (2.3.1). In particular, whenever the null hypothesis is not rejected in the latter, it will also not be rejected in the performance-based test. However, the new notion of a null hypothesis can hold under a much broader set of scenarios; the example provided in the introduction in Figure 2.1 provides such an illustration.

The new notion of the null hypothesis in (2.4.2) is attempting to capture the main question facing the decision-maker: does restricting the class of possible models to
\( L(\Theta) \) imply a significant deterioration in performance?

The reader would have obviously noted that the best achievable performance \( \pi^* \), as well as the best performance achieved by pricing based on the parametric model class \( \pi(x^*(\theta^*)) \) are not directly computable, as \( \lambda(\cdot) \) is not known to the decision-maker. The remaining challenge is therefore to prescribe a procedure for executing the performance-based test so as to meet a required significance level as well as the requirement of consistency. This is spelled out in the subsequent section.

### 2.5 The Performance-Based Approach: Proposed Test

We now turn to the analysis of the performance-based hypothesis test presented in (2.4.2)-(2.4.3) and in particular to the question of designing an appropriate test statistic based on which one can determine whether to reject \( H_0 \) or not based on the observed data of consumer purchasing decisions.

#### 2.5.1 The nonparametric approach

The first step towards operationalizing the test (2.4.2)-(2.4.3) is to define a consistent estimator of the true profit function \( \pi(x) \). Note that the postulated model class \( L(\Theta) \) need not contain the response function \( \lambda(\cdot) \). Letting \( Z_i = r(X_i)Y_i \) for \( i = 1, \ldots, n \), the available data can be viewed as noisy observations of the profit rate at the \( n \) discrete points \( X_i \)'s. Indeed,

\[
Z_i = r(X_i)Y_i = \pi(X_i) + \varepsilon(X_i),
\]  

(2.5.1)
where \( \varepsilon(X_i) = r(X_i)Y_i - \pi(X_i) \), and given \( X_i \), \( \varepsilon(X_i) \) is a random variable with zero mean and variance \( \sigma^2(X_i) := (r(X_i))^2 \lambda(X_i)(1 - \lambda(X_i)) \). There are various nonparametric approaches to estimating \( \pi(\cdot) \) (cf. Härdle (1990)). One of the most straightforward is the Nadaraya-Watson estimator:

\[
\hat{\pi}_n(x) := \frac{1}{nh} \sum_{i=1}^n K\left( \frac{x - X_i}{h} \right) Z_i \ll 2.5.2 \]

where \( h \), the so-called bandwidth, is a positive tuning parameter and the Kernel, \( K : \mathbb{R} \rightarrow \mathbb{R}_+ \) is such that \( \int K(u)du = 1 \) and \( K(u) = K(-u) \). Note that \( \pi(x) = \mathbb{E}[Z_i(x)|x] \) can be rewritten as \( \mathbb{E}[Z_i 1\{X_i = x\}]/g(x) \). Viewed this way, the numerator approximates the expectation \( \mathbb{E}[Z_i 1\{X_i = r\}] \), while the denominator in (2.5.2) approximates the density of \( X \) at the point \( x \), \( g(x) \). For purposes of concreteness, we will assume throughout that the Kernel is Gaussian, i.e.,

\( K(x) = (2\pi)^{-1/2} \exp\{-x^2/2\} \) for \( x \in \mathbb{R} \), noting that other choices are possible; cf. Härdle (1990).

### 2.5.2 The test statistic

Note that \( \hat{\pi}_n(\cdot) \) is continuous and let \( \hat{x}_n \in \arg\max\{\hat{\pi}_n(x) : x \in [x, \bar{x}]\} \) where \( \hat{\pi}_n(\cdot) \) is defined in (2.5.2). Recall that \( x^*(\theta) \in \arg\max\{p(x; \theta) : x \in [x, \bar{x}]\} \) and that \( \hat{\theta} \) denotes the maximum likelihood estimator of the parameter vector \( \theta \) based on the observations \( \{(X_i, Y_i) : i = 1, \ldots, n\} \). The performance-based test statistic \( \Delta_n \) is then defined as

\[
\Delta_n = \hat{\pi}_n(\hat{x}_n) - \hat{\pi}_n(x^*(\hat{\theta})). \ll 2.5.3 \]
The motivation behind this construction is as follows. As the sample size grows large, the approximation $\hat{\pi}_n(\cdot)$ should eventually provide a good approximation for the true profit function $\pi(\cdot)$, and hence we anticipate that $\hat{\pi}_n(x_n) \approx \pi(x^*)$. Similarly, $\hat{\theta}$ should be close to $\theta^*$ and this should imply that $\hat{\pi}_n(x^*(\hat{\theta})) \approx \pi(x^*(\theta^*))$.

As a result, $\Delta_n$ can be viewed as a noisy version of the difference between the two terms in the null hypothesis (2.4.2), i.e., $\Delta_n \approx \pi(x^*) - \pi(x^*(\theta^*))$. Based on this, we anticipate that $\Delta_n$, properly scaled, would converge in distribution to some random variable under $H_0$, while it would diverge under the alternative hypothesis.

To formalize this intuition we impose the following technical assumptions.

**Assumption 3 (Interior maximum)**

i.) $\pi(\cdot)$ is twice continuously differentiable $[x, \bar{x}]$ with unique maximizer $x^*$ which is interior and such that $\pi''(x^*) < 0$.

ii.) $p(\cdot; \cdot)$ is twice continuously differentiable on $[\bar{x}, \bar{x}] \times \Theta$. For all $\theta \in \Theta$, $p(\cdot; \theta)$ has unique maximizer $x^*(\theta)$ which is interior and such that $\partial^2 p(x^*(\theta); \theta)/\partial x^2 < 0$.

**Assumption 4 (Maximum Likelihood)**

i.) $\mathbb{E}[|\log \mathbb{P}(Y|X)|] < \infty$ and $|\log \mathbb{P}_{\theta}(y|x)| \leq f_0(x)$ for all $\theta \in \Theta$, where $f_0(\cdot)$ is bounded on $X$.

ii.) The KLIC $\mathbb{E}\left[\log(\mathbb{P}(Y_i|X_i)/\mathbb{P}_{\theta}(Y_i|X_i))\right]$ admits a unique minimum $\theta^* \in \Theta$.

iii.) $\log \mathbb{P}_{\theta}(y|x)$ is thrice differentiable with respect to $\theta$ on $\Theta$. 
iv.) There exist functions $f_1(\cdot)$, $f_2(\cdot)$ and $f_3(\cdot)$ bounded on $\mathcal{X}$ such that for all $\theta \in \Theta$, and for all $x, y \in \mathcal{X} \times \{0, 1\}$,

$$
\left| \frac{\partial \ell(x; \theta)}{\partial \theta_i} \right| \leq f_1(x), \quad \left| \frac{\partial^2 \ell(x; \theta)}{\partial \theta_i \partial \theta_j} \right| \leq f_2(x), \quad \left| \frac{\partial^3 \log \mathbb{P}_\theta(y|x)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| \leq f_3(x)
$$

v.) For each $\theta \in \Theta$, let $I_\theta$ be the fisher information matrix whose elements are defined as

$$
I_\theta(i, j) := \mathbb{E}_\theta \left[ \frac{\partial \log \mathbb{P}_\theta(y|x)}{\partial \theta_i} \frac{\partial \log \mathbb{P}_\theta(y|x)}{\partial \theta_j} \right] \quad i, j = 1, \ldots, d.
$$

We assume that $I_\theta$ is a positive definite matrix for all $\theta \in \Theta$.

Assumption 3 just ensures that the optimal decision is interior and Assumption 4 is standard in the context of asymptotic analysis of maximum likelihood estimators (cf. Serfling (2002)).

**Theorem 2 (Consistency Test)** Let Assumptions 1, 3 and 4 hold. Put

$$
\gamma := -\frac{1}{2\pi''(x^*)} \frac{\sigma^2(x^*)}{g(x^*)} \left( K^{(1)}(\psi) \right)^2 d\psi,
$$

and let $h_n \downarrow 0$ be a sequence of positive real numbers. If $nh_n^6 \to \infty$ and $nh_n^5 \to 0$, then

i.) Under $H_0$: $nh_n^3 \Delta_n \Rightarrow \gamma \chi^2$ as $n \to \infty$, 

ii.) Under $H_1$: $nh_n^3 \Delta_n \Rightarrow \infty$ as $n \to \infty$, 

where $\chi^2$ is a Chi-squared random variable with one degree of freedom.

**Discussion.** 1. $nh_n^3 \Delta_n$ converges under the null hypothesis to a scaled $\chi^2$ random variable, and diverges to infinity under the alternative. Given the value of the
constant $\gamma$, a consistent decision rule can be constructed on the basis of the limiting distribution under $H_0$: At a significance level $\alpha$, the decision would be to reject $H_0$ if $nh_n^3 \Delta_n > \tau_\alpha$ where $\tau_\alpha$ is the $(1 - \alpha)$ quantile of $\gamma \chi^2$. Note that the test is consistent as the probability of not rejecting $H_1$ when it is true is given by $\mathbb{P}(nh_n^3 \Delta_n \leq \tau_\alpha | H_1)$ which converges to zero as $n \to \infty$ by Theorem 2 ii.). 2. At an intuitive level, the test procedure can only distinguish the best performance achieved by the model-based decision and the best achievable performance, up to noise of order $(nh_n^3)^{-1}$. In other words, if $\Delta = \pi(x^*) - \pi(x^*(\theta^*))$ is of this order, it would be difficult to "see" any ill effects stemming from the use of the the restricted model class $\mathcal{L}(\Theta)$. 3. The value of the constant $\gamma$ plays a crucial role in the proposed test: the higher this value is, the harder it will be to reject the model. $\gamma$ is increasing in the the variance of the noise associated with observations at the optimal operating point $\sigma^2(x^*)$. This is consistent with intuition, as higher variance in the noise induces larger confidence bands around any nonparametric estimator of the profit function, and as a result makes it harder to reject any given model. Note that the "flatter" the profit function will be in the region of the optimum (i.e., the smaller the value of $|\pi''(x^*)|$), the harder it will be to reject a model. This stems from the fact that operating away from the optimum will have lesser ill effects on performance. This is highlighted by the fact that $\gamma$ is inversely proportional to $\pi''(x^*)$.

**An implementable test.** As the above discussion makes clear, the value of the constant $\gamma$ depends on characteristics of the true profit function, which is not
known to the decision-maker. Consequently, the asymptotic result in Theorem 2 is not directly implementable. What is needed is an approximation of $\gamma$. Indeed, $g(x^*)$ can be approximated by $\hat{g}(\hat{x}_n)$, where $\hat{g}(x)$ is given in the denominator of (2.5.2). Then $\pi''(x^*)$ can be approximated by $\hat{\pi}''(\hat{x}_n)$, and $\sigma^2(x^*)$ can be approximated by using a Kernel approximation to compute $\mathbb{E}[(Z_i - E[Z_i])^2|X_i]$ as follows

$$\hat{\sigma}^2(\hat{x}_n) = \frac{1}{n} \sum_{i=1}^{n} \frac{K\left(\frac{x-X_i}{h}\right)(Z_i - \hat{\pi}(\hat{x}_n))^2}{\frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)}.$$ 

(2.5.5)

Letting

$$\hat{\gamma}_n := -\frac{1}{\min\{2\sigma^2(\hat{x}_n)\hat{g}(\hat{x}_n); -1/n\}} \hat{\sigma}^2(\hat{x}_n) \int (K^{(1)}(\psi))^2 d\psi + \frac{1}{n},$$

(2.5.6)

we have that $\hat{\gamma}_n$ converges to $\gamma$ in probability; the corrections by $1/n$ are just introduced to ensure that the denominator is different than zero and that $\hat{\gamma}_n > 0$.

The following corollary provides an implementable version of the performance-based test

**Theorem 3 (Implementable test)** Let Assumptions 1, 3 and 4 hold. If $h_n = c_h n^{-1/7} / (\log n)^{1/7}$ where $c_h$ is a positive constant, then

i.) Under $H_0$: $n^{4/7}(\log n)^{-3/7} \hat{\gamma}_n^{-1} \Delta_n \Rightarrow \chi^2$ as $n \to \infty$,

ii.) Under $H_1$: $n^{4/7}(\log n)^{-3/7} \hat{\gamma}_n^{-1} \Delta_n \Rightarrow \infty$ as $n \to \infty$.

We now have an implementable test that allows one to evaluate if the profits achieved using prices based on the postulated parametric demand model differ in a statistically significant manner from those that would be realized using the optimal decision based on the true model. It is worth stressing that failure to
reject the null does not imply that we have identified the “right” model, nor that we have used the right decision. Rather it asserts that based on the available data, one cannot distinguish the performance induced by the decision based on the parametric model and the best achievable performance, had we known the true underlying model.

**An alternative procedure to estimate $\gamma$.** The weak convergence result in Theorem 3 i.) depends critically on the rate at which $\hat{\gamma}_n$ converges to $\gamma$. Given that this involves estimating the second derivative of the profit function, we expect this convergence to be slow. To improve the finite sample performance of the test, we propose a bootstrapping procedure for estimating the constant $\gamma$.

Algorithm 1: Bootstrapping estimate of $\gamma$

1) For a fixed positive integer $b$ and $j = 1, ..., b$, draw $n$ vectors with replacements from $\{(X_i, Y_i) : 1 \leq i \leq n\}$. Let $D_n^{(j)} = \{(X_i^{(j)}, Y_i^{(j)}) : 1 \leq i \leq n\}$ denote the resulting draw, and let $\hat{\pi}_n^{(j)}(\cdot)$ denote the Kernel based estimator of the profit function based on the dataset $D_n^{(j)}$.

2) Let $\hat{x}_n^{(j)} \in \arg\max\{\hat{\pi}_n^{(j)}(x) : x \in [\bar{x}, \bar{x}]\}$ and

$$
\Delta_n^{(j)} = \hat{\pi}_n^{(j)}(\hat{x}_n^{(j)}) - \hat{\pi}_n^{(j)}(\hat{x}_n).
$$

(2.5.7)

3) Let

$$
\hat{\gamma}_n^b = \frac{1}{b} \sum_{j=1}^{b} \Delta_n^{(j)}. \tag{2.5.8}
$$
4) For the significance level $\alpha$, let $\tau_\alpha$ to be the $(1 - \alpha)$ quantile of $\hat{\gamma}_n^2$. Then one rejects the null if and only if $nh_n^2 \Delta_n > \tau_\alpha$.

Based on classical results on bootstrapping (see, e.g., Efron and Tibshirani (1993) and Giné and Zinn (1990)) we expect that $\hat{\gamma}_n^2$ converges to $\gamma$ under $H_0$ as $n$ and $n_b$ grow to $\infty$. While spelling out this limit theory is beyond the scope of this paper, we compare the results obtained using the bootstrapping procedure with those using the estimation procedure (2.5.6) in the next section.

### 2.6 Properties of the Proposed Test: Illustrative Numerical Examples

We present below numerical results that illustrate the properties of the performance-based test, and contrast those with the model-based test discussed in Section 2.3.2.

For illustrative purposes, we consider two response function models: a logit structure with parameters $(\theta_1, \theta_2)$, $\ell(x; \theta) = \exp\{\theta_1 - \theta_2 x\}(1 + \exp\{\theta_1 - \theta_2 x\})^{-1}$; and an exponential structure with parameter $(\theta_1, \theta_2)$, $\ell(x; \theta) = \theta_1 \exp\{-\theta_2 x\}$. In all cases, the revenue function is given by $r(x) = (x - 1)$, and the $X_i$'s are drawn form a uniform distribution on $[1, 9]$. The bandwidth is taken to be $h_n = c_h n^{-1/7}/(\log n)^{1/7}$, with $c_h > 0$, a tuning constant whose effects are examined below. The procedure we follow is to approximate the distribution of the scaled test statistic on the basis of Theorem 3, or based on the bootstrapping procedure that was discussed above.

Given the limit distribution and a significance level $\alpha$, we define the rejection region for the null. Say, for an estimated value of the scaling constant $\hat{\gamma}$, we take the $(1 - \alpha)$ quantile $\tau_\alpha$ such that $\mathbb{P}\{\hat{\gamma}^2 \chi^2 > \tau_\alpha\} = \alpha$. In our experiments, we focus
on a rather standard choice of $\alpha = 5\%$.

In Table 2.1, we consider a scenario where the data is generated according to a logit with parameters $(\theta_1 = 3, \theta_2 = -.9)$, and the assumed structure is well-specified: in this case, the null hypothesis will be true in both the model-based formulation (2.3.1) as well as the performance-based one (2.4.2). We present the impact of the constant $c_h$, i.e., of the bandwidth choice and compare the accuracy of the test using the asymptotic distribution in Theorem 3, versus the bootstrap procedure in Algorithm 1. In particular, we depict the number of times one rejects the null at the 5% level based on 500 replications.

<table>
<thead>
<tr>
<th>$c_h$</th>
<th>1.5</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>performance-based test</td>
<td>18.2%</td>
<td>8.4%</td>
<td>4.0%</td>
<td>0.8%</td>
</tr>
<tr>
<td>performance-based test (boot)</td>
<td>9.8%</td>
<td>8.2%</td>
<td>4.4%</td>
<td>1.2%</td>
</tr>
</tbody>
</table>

Table 2.1: **Efficacy of the performance-based test.** Fraction of time one rejects $H_0$ at the 5% level (based on 500 replications) and the effect of the bandwidth parameter. The data-generating model is a logit and the assumed structure is a logit (well-specified case).

It is evident that the choice of bandwidth parameter $c_h$ impacts the behavior of the test statistic and it seems that a constant $c_h$ in the range $[2, 3]$ is appropriate. We note that the bootstrapping procedure provides more consistent results across bandwidths, and improves the finite sample performance of the test.

In Table 2.2, we focus again on the fraction of time one rejects the null at the 5% level based on 500 replications. We compare the results provided by the performance-based test to those of the model-based test. In the first (case 1), the true model is a logit with parameters $(\theta_1 = 3, \theta_2 = -.9)$ and the assumed structure
is also logit. In other words, the true response function belongs to the postulated family and \( H_0 \) holds true in this case from both model-based and performance-based perspectives. The second (case 2) considers again a true model which is logit with parameters as above but the assumed structure is exponential \( \ell(x; \theta) = \theta_1 \exp\{-\theta_2 x\}. \) In the last (case 3), the true model is a logit with parameters \( (\theta_1 = 4.5, \theta_2 = -0.9) \) and the assumed structure is exponential. The bandwidth for the performance-based test is taken to be \( h_n = 2n^{-1/7}/(\log n)^{1/7}. \)

<table>
<thead>
<tr>
<th>true model</th>
<th>case 1</th>
<th>case 2</th>
<th>case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>logit: (3, -0.9)</td>
<td>logit: (3, -0.9)</td>
<td>logit: (4.5, -0.9)</td>
<td></td>
</tr>
<tr>
<td>assumed structure</td>
<td>logit: ( (\theta_1, \theta_2) )</td>
<td>exp: ( (\theta_1, \theta_2) )</td>
<td>exp: ( (\theta_1, \theta_2) )</td>
</tr>
<tr>
<td>Data size ( (n) )</td>
<td>( 5 \times 10^2 )</td>
<td>( 10^3 )</td>
<td>( 5 \times 10^3 )</td>
</tr>
<tr>
<td>model-based test</td>
<td>3.0%</td>
<td>6.2%</td>
<td>98.6%</td>
</tr>
<tr>
<td>performance-based test</td>
<td>8.2%</td>
<td>8.4%</td>
<td>12.8%</td>
</tr>
</tbody>
</table>

Table 2.2: **Comparison of the performance-based test and model-based test.** Fraction of time one rejects \( H_0 \) at the 5% level (based on 500 replications).

Focusing exclusively on the performance-based test, we observe that when the model is well specified (case 1) and hence \( H_0 \) is correct for both model- and performance-based tests, the latter rejects \( H_0 \) about 8% of the times at the 5% level. Turning to case 2, where the assumed structure is incorrect, we observe that the model-based test rejects the exponential model more than 98% of the times. This is in sharp contrast with the performance-based test that only rejects the exponential model about 13% of the times. In case 3, where again the assumed structure is misspecified relative to the true response function, both tests reject the null more than 90% of the times. To better understand the phenomena at play in cases 2 and 3, we present in Figures 2.2 and 2.3 the true demand model (logit)
and the "best" exponential fit (see discussion in Section 2.4). We observe that the discrepancy between the two response curves are quite noticeable, exceeding in places 10% in absolute value (Figures 2.2(a), 2.3(a)). Analyzing the profit curves corresponding to case 2 in Figure 2.2(b), we observe that the difference between the optimal performance under the true model, and the performance of the decision dictated by the best exponential fit differ by an amount $\Delta$ which is quite minute. Thus the performance-based test indicates that it is difficult to distinguish the difference in performance with dataset sizes of 500 or 1000. In case 3, focusing on the profit curves in Figure 2.3(b), we see that the difference $\Delta$ becomes significant when compared to case 2. This is why the performance-based test now rejects the null more than 90% of the time. Note also that the amount of times one rejects the null increases with the size of the data, illustrating that the difference $\Delta$ becomes more significant for larger sample sizes.

2.7 Empirical Example

This section focuses on applying the proposed performance-based test to the case of an auto lender operating in the online direct-to-consumer sales channel. We provide a description of the dataset in Section 2.7.1, then discuss the setup and objective of the firm in Section 2.7.2 and present the results of the performance-based test in Section 2.7.3.
Figure 2.2: **Model misspecification that is not rejected by the performance-based test.** The true model is a logit with parameters ($\theta_1 = 3, \theta_2 = -0.9$). Panel (a) gives the true demand model and the best exponential fit; Panel (b) depicts the true profit function and that based on the best exponential fit. $\Delta$ indicates the difference between the optimal performance, and the one achieved by the optimal decision based on the best exponential fit.

### 2.7.1 Data description

The loan process can be described as follows: i.) first, a customer fills an online loan request consisting of the amount and term desired as well as a questionnaire with personal data; ii.) the lender then analyzes the loan application and either directly discards the application or quotes a rate to the applicant; iii.) the customer, upon receiving the offer, has a few weeks to decide whether to accept the offer or not.

The data contains all instances of incoming customers who are ultimately offered a loan during the period ranging from July 2002 to December 2004. For each such customer $i$, the following information was available:
Figure 2.3: Model misspecification that is not rejected by the performance-based test. The true model is a logit with parameters \( \theta_1 = 4.5, \theta_2 = -0.9 \). Panel (a) gives the true demand model and the best exponential fit; Panel (b) depicts the true profit function and that based on the best exponential fit. \( \Delta \) indicates the difference between the optimal performance, and the one achieved by the optimal decision based on the best exponential fit.

1) characteristics of the loan requested:

a) date of the request

b) amount requested (in dollars), denoted \( W_{1,i} \).

c) term requested (in months), denoted \( W_{2,i} \).

d) loan type, denoted \( W_{3,i} \). This variable indicates whether the loan was used for a used car, a new car or to refinance a car.

2) annual percentage rate that was quoted to the customer; this was decided by the firm.
3) decision of the customer:

a) accept/reject decision, denote $Y_i$. This is a binary variable indicating whether the customer accepted or rejected the offer.

b) date of acceptance of the offer for customers who accepted the rate quoted.

4) customer characteristics:

a) the FICO score, denoted $W_{4,i}$. FICO is computed through a proprietary algorithm using the credit history of customers and has become a standard measure to quantify the risk associated with lending to a given customer. The score ranges from 300 to 850 and the higher the score the lower the probability of default of a customer.

b) the state in which the customer lives, denoted $W_{5,i}$.

2.7.2 Problem definition

Let $W$ be the vector summarizing the loan/customer characteristics. The firm would ideally like to have a handle on the response function, $\lambda_W(x)$, i.e., the probability of acceptance as a function of the quoted rate $x$ for a given loan/customer profile. If $x_0$ denotes the cost of funds faced by the lending firm, the profit maximization problem can be approximated by

$$\max_{r \geq x_0} (x - x_0 - \text{risk factor})\lambda_W(x),$$

(2.7.1)

where the risk factor might also depend on other characteristics of the loan and the customer. The risk factor term was not available and we will not consider it.
for this illustration. However, it should be apparent that a proper description of
the risk factor, if available, can easily be incorporated. In what follows, we will
take $x_0 = 2\%$. 

Rather than solving (2.7.1) for every profile $W$, we will segment the space of
profiles $W$ along various dimensions and attempt to maximize the profits on
each segment. For the purposes of this study, we will focus on the data from
incoming customers in the period from December 2003 to December 2004, loans
for used cars only and customers with FICO scores in the range 690 to 740. Within
that group of customers/loans that we focus on, we segment customers with two
possible FICO score ranges $((690—715]$ and $(715—740]$) and four possible requested
terms: term 1, term 2, term 3, term 4.

For each segment, the parametric family of models for the acceptance proba-

\[ \ell(x; \theta_1, \theta_2) = \frac{\exp\{\theta_1 + \theta_2 x\}}{1 + \exp\{\theta_1 + \theta_2 x\}}, \]  

(2.7.2)

and $\theta_1$ and $\theta_2$ are two parameters.

2.7.3 Results

We present in Table 2.3 the results we obtain when applying the model-based
test and the performance-based test to four of the segments described earlier over
the period of a year. The four other segments either did not have an interior
maximum or had insufficient data. For the experiments, 250 bootstrap samples
were used for both tests and the bandwidth $h$ for the performance-based test was
\[ h_n = (2.25)n^{-1/7}/(\log n)^{1/7}. \]

<table>
<thead>
<tr>
<th></th>
<th>term 1</th>
<th>term 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>range 1</strong> (690-715)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sample size</td>
<td>n</td>
<td>592</td>
</tr>
<tr>
<td>parameter estimates</td>
<td>( \hat{\theta}_1, \hat{\theta}_2 )</td>
<td>(1.15, -0.35)</td>
</tr>
<tr>
<td>model-based test p-value</td>
<td>8.8%</td>
<td>0.4%</td>
</tr>
<tr>
<td>performance-based test p-value</td>
<td>22.5%</td>
<td>1.8%</td>
</tr>
<tr>
<td><strong>range 2</strong> (715-740)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sample size</td>
<td>n</td>
<td>658</td>
</tr>
<tr>
<td>parameter estimates</td>
<td>( \hat{\theta}_1, \hat{\theta}_2 )</td>
<td>(1.59, -0.41)</td>
</tr>
<tr>
<td>model-based test p-value</td>
<td>0.0%</td>
<td>9.2%</td>
</tr>
<tr>
<td>performance-based test p-value</td>
<td>58.6%</td>
<td>81.5%</td>
</tr>
</tbody>
</table>

Table 2.3: Comparison of the model-based test and the performance-based test for the Logit model (250 bootstraps).

We observe that at the 5% level, one rejects the logit model in two out four segments. In contrast, the performance-based rejects the logit model only in one of them. At the 10% level, one rejects the logit model in all four segments based on the model-based test while the logit is rejected in only one segment in one of them. In other words, the logit model appears to be "good enough" for purposes of revenue management in three out of four segments eventhough the logit is not necessarily a good global fit. In order to illustrate in further detail this point, we depict in Figure 2.4 the logit fit and the nonparametric estimates of both the response and profit functions for the segment with term 3 and FICO scores in (715 – 740]. We observe that while the response functions differ for example for rates above 6.5%, the performances associated with the decisions dictated by the best logit fit and the nonparametric fit are indistinguishable, leading to a very high p-value for the performance based-test.
Figure 2.4: Model misspecification that is not rejected by the performance-based test. Illustration of the result in the segment with FICO range 715-740 and Term 3. \( \hat{x} \) is the maximizer of the Kernel based approximation to the profit function and \( x^*(\hat{\theta}) \) is the maximizer of the profit function based on the best logit fit. Here, \( \Delta_n \) is too small to be indicated on the figure.
Chapter 3

The Dynamic Pricing Problem under Model Uncertainty

3.1 Introduction

A critical assumption made in most academic studies of revenue management problems is that the functional relationship between the mean demand rate and price, often referred to as the demand function or demand curve, is known to the decision maker. This makes the underlying problems more tractable and allows one to extract structural insights. At the same time, this assumption of "full information" endows the decision maker with knowledge that s/he does not typically possess in practice.

Lack of information concerning the demand model raises several fundamental questions. First, and foremost, is it possible to quantify the "value" of full information (for example, by measuring the revenue loss due to imperfect information)? Second, is it possible to achieve anything close to the maximal revenues in the full information setting by judiciously combining real-time demand learning and pricing strategies? Finally, how would such strategies exploit prior information, if
any, on the structure of the demand function?

The main objective of this chapter is to shed some light on the aforementioned questions. Our departure point will be a prototypical single product revenue management problem first introduced and formalized by Gallego and van Ryzin (1994). This formulation models realized demand as a Poisson process whose intensity at each point in time is determined by a price set by the decision maker. Given an initial inventory, the objective is to dynamically price the product so as to maximize expected revenues over a finite selling horizon. In the dynamic optimization problem considered in Gallego and van Ryzin (1994), the decision maker knows the demand function prior to the start of the selling season and designs optimal policies based on this information. In the setting we pursue in this chapter it is only possible to observe realized demand over time, and the demand function itself is not known. To that end, we consider two levels of uncertainty with regard to the demand model: i.) a nonparametric setting where the demand function is only assumed to belong to a broad functional class satisfying mild regularity conditions; and ii.) a parametric setting in which the demand function admits a given parametric structure but the parameter values are not known.

The absence of perfect prior information concerning the demand model introduces an important new component into the above dynamic optimization problem, namely, tension between exploration (demand learning) and exploitation (pricing). The longer one spends learning the demand characteristics, the less time remains to exploit that knowledge and optimize profits. On the other hand, less time spent on
demand learning leaves more residual uncertainty that could hamper any pricing strategy in the exploitation phase. One of the main contributions of this chapter is to formulate this dynamic pricing problem under incomplete information, and to pursue an analysis that highlights the key tradeoffs discussed above, and articulate them in a precise mathematical manner.

To address uncertainty with regard to the demand model, we introduce a family of pricing policies that learn the demand function "on the fly." Their performance will be measured in terms of the revenue loss relative to a full information benchmark that assumes knowledge of the demand function. We refer to this loss as the regret associated with not knowing the demand function a priori; the magnitude of the regret quantifies the economic value of prior model information. The policies we consider are designed with the objective of achieving a "small" regret uniformly over the relevant class of demand functions (either parametric or nonparametric). This adversarial setting, where nature is allowed to counter a chosen policy with the "worst" demand function, ensures that policies exhibit "good" performance irrespective of the true demand model.

The complexity of the problem described above makes it difficult to evaluate the performance of any reasonable policy, except via numerical experiments. To address this issue, we consider an asymptotic regime which is characterized by a high volume of sales. More specifically, the initial level of inventory and the magnitude of demand ("market size") grow large in proportion to each other; see Gallego and van Ryzin (1994) and Talluri and van Ryzin (2005) for further exam-
amples in the revenue management literature that adopt this framework. This regime allows us to bound the magnitude of the regret, and to establish a rather surprising result with regard to our proposed policies: as the sales volume grows large, the regret eventually shrinks to zero. That is, these policies achieve (asymptotically) the maximal full information revenues, despite the absence of prior information regarding the demand function; in that sense, they are asymptotically optimal.

In more detail, the main contributions of this chapter are summarized as follows.

i.) We introduce a nonparametric pricing policy (see Algorithm 1) that requires almost no prior information on the demand function. In settings where the structure of the demand function is known up to the value/s of certain parameter/s, we develop a parametric pricing policy based on Maximum Likelihood estimation (see Algorithm 2 and Algorithm 3).

ii.) We establish lower bounds on the regret that hold for any admissible learning and pricing policy (see Propositions 2 and 4).

iii.) We derive upper bounds on the performance of our nonparametric and parametric pricing policies (see Propositions 1 and 3). In all cases the proposed policies achieve a regret that is “not far” from the lower bound described above. In the parametric setting when only one parameter is unknown, we prove that essentially no admissible pricing policy can achieve a smaller regret than our proposed method (see Proposition 5).

iv.) Building on ideas from stochastic approximations, we indicate how one can
develop more refined sequential policies, and illustrate this in the nonpara-
metric setting (see Algorithm 4).

Returning to the questions raised earlier in this section, our results shed light on the following issues. First, despite having only limited (or almost no) prior information, it is possible to construct joint learning and pricing policies that generate revenues which are “close” to the best achievable performance with full information. Our results highlight an interesting observation. In the full information setting, Gallego and van Ryzin (1994) prove that fixed price heuristics lead to near-optimal revenues, hence the value of dynamic price changes is (at least asymptotically) limited. In a setting with incomplete information, price changes play a much more pivotal role, as they are relied upon to resolve uncertainty with regard to the demand function.

The regret bounds described above rigorously quantify the economic value of a priori information on the demand model. Alternatively, the lost revenues can be viewed as quantifying the “price” paid for model uncertainty. Finally, our work highlights an important issue related to model misspecification risk. In particular, if an algorithm is designed under parametric assumptions, it is prone to such risk as the true demand function may not (and in many cases will not) belong to the assumed parametric family. Our regret bounds provide a means for quantifying the “price” that one pays for eliminating this risk via nonparametric approaches; see also the numerical illustration in Section 3.6.

The remainder of the chapter. The next section reviews related literature.
Section 3.3 introduces the model and formulates the problem. Section 3.4 studies the nonparametric setting and Section 3.5 focuses on cases where the demand function possesses a parametric structure. Section 3.6 presents numerical results and discusses some qualitative insights. Section 3.7 formulates adaptive versions of the nonparametric algorithm and illustrates their performance. All proofs are collected in two appendices: Appendix B.1 contains the proofs of the main results; and Appendix B.2 contains proofs of auxiliary lemmas.

3.2 Related Literature

Parametric approaches. The majority of revenue management studies that address demand function uncertainty do so by assuming that one or more parameters characterizing this function are unknown. The typical approach here follows a dynamic programming formulation with Bayesian updating, where a prior on the distribution of the unknown parameters is initially postulated. Recent examples include Aviv and Pazgal (2005), Araman and Caldentey (2005) and Farias and Van Roy (2006), all of which assume a single parameter is unknown. (See also Lobo and Boyd (2003) and Carvalho and Puterman (2005).) Scarf (1959) was one of the first papers to use this Bayesian formulation, though in the context of inventory management.

While the Bayesian approach provides for an attractive stylized analysis of the joint learning and pricing problem, it suffers from significant shortcomings. Most notably, the objective of the dynamic optimization problem involves an expecta-
tion that is taken relative to a prior distribution over the unknown parameters. Hence any notion of optimality associated with a Bayesian-based policy is with respect to that prior. Moreover, the specification of this prior distribution is typically constrained to so-called conjugate families and is not driven by "real" prior information; the hindering element here is the computation of the posterior via Bayes rule. The above factors introduce significant restrictions on the models that are amenable to analysis via Bayesian dynamic programming. Bertsimas and Perakis (2003) have considered an alternative to this formulation using a least squares approach in the context of a linear demand model. Our work in the parametric setting is based on maximum likelihood estimation, and hence is applicable to a wide class of parametric models; as such, it offers an alternative to Bayesian approaches that circumvents some of their deficiencies.

Nonparametric approaches. The main difficulty facing nonparametric approaches is loss of tractability. Most work here has been pursued in relatively simple static settings that do not allow for learning of the demand function; see, e.g., Ball and Queyranne (2006) and Eren and Maglaras (2007) for a competitive ratio formulation, and Perakis and Roels (2006) for a minimax regret formulation. These studies focus almost exclusively on structural insights. The recent paper by Rusmevichientong et al. (2006) develops a nonparametric approach to a multiproduct static pricing problem, based on historical data. The formulation does not incorporate inventory constraints or finite sales horizon considerations. In the context of optimizing seat allocation policies for a single flight multi-class prob-
lem, van Ryzin and McGill (2000) show that one can use stochastic approximation methods to reach near-optimal capacity protection levels in the long run. The absence of parametric assumptions in this case is with respect to the distributions of customers' requests for each class. (See also Huh and Rusmevichientong (2006) for a related study.)

Perhaps the most closely related paper to our current work is that of Lim and Shanthikumar (2006) who formulate a robust counterpart to the single product revenue management problem of Gallego and van Ryzin (1994). In that paper the uncertainty arises at the level of the point process distribution characterizing realized demand, and the authors use a max-min formulation where nature is adversarial at every point in time. This type of conservative setting effectively precludes any real-time learning, and moreover does not lend itself to prescriptive solutions.

Related work in other disciplines. The general problem of dynamic optimization with limited or no information about a response function has also attracted attention in other fields. In economics, a line of work that traces back to Hannan (1957) studies settings where the decision maker faces an oblivious opponent. The objective is to minimize the difference between the rewards accumulated by a given policy, and the rewards accumulated by the best possible single action had the decision maker known in advance the actions of the adversary; see Foster and Vohra (1999) for a review of this line of work and its relation to developments in other fields.
A classical formulation of sequential optimization under uncertainty that captures the essence of the exploration-exploitation tradeoff, is the multiarmed bandit paradigm that dates back to the work of Robbins (1952). This was originally introduced as a model of clinical trials in the statistics literature, but has since been used in many other settings; see, e.g., Lai and Robbins (1985) and references therein. Related studies in the computer science literature include Auer et al. (2002) who study an adversarial version of a multi-armed bandit problem, and Kleinberg and Leighton (2003) who provide an analysis of an on-line posted-price auction using these tools. (See Cesa-Bianchi and Lugosi (2006) for a recent and comprehensive survey). Our work shares an important common theme with the streams of literature survey above, insofar as it too highlights exploration-exploitation tradeoffs. On the other hand, our work represents a significant departure from antecedent literature along three important dimensions that are characteristic of our dynamic pricing problem: we deal with a constrained dynamic optimization problem (the constraint arising from the initial inventory level); the action space of the decision maker, namely the feasible price set, is uncountable; and the action space of the adversary (nature), namely, the class of admissible demand functions, is also uncountable.

3.3 Problem Formulation

Model primitives and basic assumptions. We consider a revenue management problem in which a monopolist sells a single product. The selling horizon is denoted
by $T > 0$, and after this time sales are discontinued and there is no salvage value for the remaining unsold products. Demand for the product at any time $t \in [0, T]$ is given by a Poisson process with intensity $\lambda_t$ which measures the instantaneous demand rate (in units such as number of products requested per hour, say): Letting $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ denote the demand function, then if the price at time $t$ is $p(t)$, the instantaneous demand rate at time $t$ is given by $\lambda_t = \lambda(p(t))$, and realized demand is a controlled Poisson process with this intensity.

We assume that the set of feasible prices is $[p, \bar{p}] \cup p_{\infty}$, where $0 < p < \bar{p} < \infty$ and $p_{\infty} > 0$ is a price that “turns off” demand (and revenue rate), i.e., $\lambda(p_{\infty}) = 0^1$. With regard to the demand function, we assume that $\lambda(\cdot)$ is non-increasing in the price $p$, has an inverse denoted by $\gamma(\cdot)$, and the revenue rate $r(\lambda) := \lambda \gamma(\lambda)$ is concave. These assumptions are quite standard in the revenue management literature resulting in the term regular affixed to demand functions satisfying these conditions; see, e.g., Talluri and van Ryzin (2005, §7).

Let $(p(t) : 0 \leq t \leq T)$ denote the price process which is assumed to have sample paths that are right continuous with left limits taking values in $[p, \bar{p}] \cup p_{\infty}$. Let $N(\cdot)$ be a unit rate Poisson process. The cumulative demand for the product up until time $t$ is then given by $D(t) := N(\int_0^t \lambda(p(s))ds)$. We say that $(p(t) : 0 \leq t \leq T)$ is non anticipating if the value of $p(t)$ at each time $t \in [0, T]$ is only allowed to depend on past prices $\{p(s) : s \in [0, t]\}$ and demand values $\{(D(s)) : s \in [0, t]\}$. (More formally, the price process is adapted to the filtration generated by the past

---

1 The case $p_{\infty} = \infty$ can be incorporated by assuming $r(\lambda(p_{\infty})) := \lim_{p \rightarrow p_{\infty}} r(\lambda(p)) = 0$ which implies that $\lim_{p \rightarrow p_{\infty}} \lambda(p) = 0$. 
values of the demand and price processes.)

**Information structure and the economic optimization problem.** We assume that the decision maker *does not know* the true demand function $\lambda$, but is able to continuously observe realized demand at all time instants starting at time 0 and up until the end of the selling horizon $T$. The only information available regarding $\lambda$ is that it belongs to a class of admissible demand functions, $\mathcal{L}$; in Section 3.4, $\mathcal{L}$ will be taken to be a *nonparametric* class of functions, and in Section 3.5 it will be restricted to a *parametric* class. Thus, the makeup of the class $\mathcal{L}$ summarizes prior information on the demand model.

We shall use $\pi$ to denote a *pricing policy*, which, roughly speaking, maps the above information structure to a non anticipating price process $(p(t) : 0 \leq t \leq T)$. With some abuse of terminology, we will use the term "policy" to refer to the price process itself and the algorithm that generates it, interchangeably. Put

$$N^\pi(t) := N\left(\int_0^t \lambda(p(s)) ds\right) \quad \text{for} \quad 0 \leq t \leq T,$$

(3.3.1)

where $N^\pi(t)$ denotes the cumulative demand up to time $t$ under the policy $\pi$.

Let $x > 0$ denote the inventory level (number of products) at the start of the selling season. A pricing policy $\pi$ is said to be *admissible* if the induced price process satisfies

$$\int_0^T dN^\pi(s) \leq x \quad \text{a.s.,}$$

(3.3.2)

$$p(s) \in [p, \overline{p}] \cup p_{\infty}, \quad 0 \leq s \leq T.$$

(3.3.3)

It is important to note that while the decision maker does not know the demand
function, knowledge that $\lambda(p_\infty) = 0$ guarantees that the constraint (3.3.2) can be met. Let $\mathcal{P}$ denote the set of admissible pricing policies.

The dynamic optimization problem faced by the decision maker under the information structure described above is: choose $\pi \in \mathcal{P}$ to maximize the total expected revenues

$$J^\pi(x, T; \lambda) := \mathbb{E}\left[\int_0^T p(s)dN^\pi(s)\right].$$

(3.3.4)

The dependence on $\lambda$ in the left-hand-side is indicative of the fact that the expectation on the right-hand-side is taken with respect to the true demand distribution. Since the decision maker cannot compute the expectation in (3.3.4) without knowing the underlying demand function, the above optimization problem does not seem to be well posed. In a sense one can view the solution of (3.3.4) as being made possible only with the aid of an “oracle” which can compute the quantity $J^\pi(x, T; \lambda)$ for any given policy. We will now redefine the decision maker’s objective in a more suitable manner, using the notion of a full information benchmark.

**A full information benchmark.** Let us first explain how the analysis of the dynamic optimization problem described in (3.3.4) proceeds when one removes two significant obstacles: lack of knowledge of the demand function $\lambda$ prior to the start of the selling season; and stochastic variability in realized demand. In particular, consider the following full information deterministic optimization problem, in
which the function $\lambda$ is assumed to be known at time $t = 0$:

$$ \sup \int_0^T r(\lambda(p(s)))ds, \quad (3.3.5) $$

s.t. \begin{align*}
\int_0^T \lambda(p(s))ds \leq x, \\
p(s) \in [p, \bar{p}] \cup \mathbb{F}_\infty \text{ for all } s \in [0, T].
\end{align*}

This problem is obtained from (3.3.4), and the admissibility conditions (3.3.2)-(3.3.3), by replacing the random process characterizing customer purchase requests by its mean rate. For example, if one focuses on the objective (3.3.4), then the deterministic objective in (3.3.5) is obtained by substituting "$\lambda(p(s))ds$" for "$dN^\pi(s)$" since $r(\lambda(p(s))) = p(s)\lambda(p(s))$. The same parallel can be drawn between the first constraint of the deterministic problem and (3.3.2). Consequently, it is reasonable to refer to (3.3.5) as a \textit{full information deterministic relaxation} of the original dynamic pricing problem (3.3.4).

Let us denote the value of (3.3.5) as $J^D(x, T; \lambda)$ where 'D' is mnemonic for deterministic and the choice of notation with respect to $\lambda$ reflects the fact that the optimization problem is solved "conditioned" on knowing the true underlying demand function. The value of the full information deterministic relaxation provides, as one would anticipate, an upper bound on expected revenues generated by any pricing policy $\pi \in \mathcal{P}$, that is, $J^*(x, T; \lambda) \leq J^D(x, T; \lambda)$ for all $\lambda \in \mathcal{L}$; this rather intuitive observation is formalized in Lemma 4 in Appendix B.1 which essentially generalizes Gallego and van Ryzin (1994, Proposition 2).

The \textbf{minimax regret objective}. As indicated above, for any demand func-
tion \( \lambda \in \mathcal{L} \), we have that \( J^\pi(x, T; \lambda) \leq J^D(x, T|\lambda) \) for all admissible policies \( \pi \in \mathcal{P} \). With this in mind, we define the regret \( R^\pi(x, T; \lambda) \), for any given function \( \lambda \in \mathcal{L} \) and policy \( \pi \in \mathcal{P} \), to be

\[
R^\pi(x, T; \lambda) = 1 - \frac{J^\pi(x, T; \lambda)}{J^D(x, T|\lambda)}.
\]  

(3.3.6)

The regret measures the percentage loss in performance of any policy \( \pi \) in relation to the benchmark \( J^D(x, T|\lambda) \). By definition, the value of the regret always lies in the interval \([0, 1]\), and the smaller the regret, the better the performance of a policy \( \pi \); in the extreme case when the regret is zero, then the policy \( \pi \) is guaranteed to extract the maximum full information revenues. Since the decision maker does not know which demand function s/he will face in the class \( \mathcal{L} \), it is attractive to design pricing policies that perform well irrespective of the actual underlying demand function. In particular, if the decision maker uses a policy \( \pi \in \mathcal{P} \), and nature then "picks" the worst possible demand function for that policy, then the resulting regret would be

\[
\sup_{\lambda \in \mathcal{L}} R^\pi(x, T; \lambda).
\]  

(3.3.7)

In this game theoretic setting it is now possible to restate the decision maker's objective, initially given in (3.3.4), as follows: pick \( \pi \in \mathcal{P} \) to minimize (3.3.7). The advantage of this formulation is that the decision maker's problem is now well posed: for any \( \pi \in \mathcal{P} \) and fixed \( \lambda \in \mathcal{L} \) it is possible, at least in theory, to compute the numerator on the right hand side in (3.3.6), and hence (3.3.7). Roughly speaking, one can attach a worst case \( \lambda \in \mathcal{L} \) to "each" policy \( \pi \in \mathcal{P} \), and subsequently one can try to "optimize" this by searching for the policy with the
best worst-case performance. In other words, we are interested in characterizing
the minimax regret
\[
\inf_{\pi \in P} \sup_{\lambda \in \mathcal{L}} \mathcal{R}^\pi(x, T; \lambda). \tag{3.3.8}
\]
This quantity has an obvious physical interpretation: it measures the monetary
value (in normalized currency units) of knowing the demand function a priori.
The issue of course is that, barring exceedingly simple cases, it is not possible to
compute the minimax regret. Our objective in what follows will be to characterize
this quantity by deriving suitable bounds on (3.3.8) in cases where the class \( \mathcal{L} \) is
nonparametric or is restricted to a suitable parametric class of demand functions.

3.4 Main Results: The Nonparametric Case

3.4.1 A nonparametric pricing algorithm

We introduce below a learning and pricing policy defined through two tuning pa‐
rameters \((\kappa, \tau)\): \(\kappa\) is a positive integer and \(\tau \in (0, T]\). The general structure,
which is summarized for convenience in algorithmic form, is divided into two main
stages. A “learning” phase (exploration) of length \(\tau\) is first used, in which \(\kappa\) prices
are tested. Then a “pricing” phase (exploitation) fixes a “good” price based on
demand observations in the first phase. The intuition underlying the method is
discussed immediately following the description of the method.

\begin{algorithm}
\begin{align*}
\text{Algorithm 1 : } & \pi(\tau, \kappa) \\
\text{Step 1. Initialization:}
\end{align*}
\end{algorithm}
(a) Set the learning interval to be $[0, \tau]$, and the number of prices to experiment with to be $\kappa$. Put $\Delta = \tau / \kappa$.

(b) Divide $[\underline{p}, \bar{p}]$ into $\kappa$ equally spaced intervals and let $\{p_i, i = 1, ..., \kappa\}$ be the left endpoints of these intervals.

**Step 2. Learning/experimentation:**

(a) On the interval $[0, \tau]$ apply $p_i$ from $t_{i-1} = (i - 1) \Delta$ to $t_i = i \Delta$, $i = 1, 2, ..., \kappa$, as long as inventory is positive. If no more units are in stock, apply $p_\infty$ up until time $T$ and STOP.

(b) Compute

$$\hat{d}(p_i) = \frac{\text{total demand over } [t_{i-1}, t_i]}{\Delta}, \quad i = 1, ..., \kappa.$$  

**Step 3. Optimization:**

Compute $\hat{p}^u = \arg \max_{1 \leq i \leq \kappa} \{p_i \hat{d}(p_i)\}$, $\hat{p}^c = \arg \min_{1 \leq i \leq \kappa} |\hat{d}(p_i) - x/T|$, \hspace{1cm} (3.4.1)

and set $\hat{p} = \max\{\hat{p}^c, \hat{p}^u\}$. \hspace{1cm} (3.4.2)

**Step 4. Pricing:**

On the interval $(\tau, T]$ apply $\hat{p}$ as long as inventory is positive, then apply $p_\infty$ for the remaining time.

---

**Intuition and key underlying ideas.** At first, a nonparametric empirical estimate of the demand function is obtained based on a learning phase of length
\(\tau\) described in Steps 1 and 2. The intuition underlying Steps 3 and 4 is based on
the analysis of the deterministic relaxation (3.3.5) whose solution (see Lemma 4 in
Appendix B.1) is given by \(p(s) = p^D := \max\{p^u, p^c\}\) for \(s \in [0, T']\) and \(p(s) = p_\infty\)
for \(s > T'\), where

\[
p^u = \arg\max_{p \in [\underline{p}, \overline{p}]} \{r(\lambda(p))\}, \quad p^c = \arg\min_{p \in [\underline{p}, \overline{p}]} |\lambda(p) - x/T|,
\]

and \(T' = \min\{T, x/\lambda(p^D)\}\). Here, the superscripts "\(u\)" and "\(c\)" stand for unconstrained and constrained, respectively, in reference to whether the inventory con-
straint in (3.3.5) is binding or not. In particular, the deterministic problem (3.3.5)
can be solved by restricting attention to the two prices \(p^u\) and \(p^c\). Algorithm 1
hinges on this observation.

In Step 3 the objective is to obtain an accurate estimate of \(p^D\) based on the
observations during the "exploration" phase, while at the same time keeping \(\tau\)
"small" in order to limit the revenue loss over this learning phase. The algorithm
then applies this price on \((\tau, T]\). With the exception of the "short" initial phase
\([0, \tau]\), the expected revenues (in the real system) will be close to those achieved
by \(p^D\) over \([0, T]\). The analysis in Gallego and van Ryzin (1994) establishes that
those revenues would be close to \(J^D(x, T|\lambda)\), and as a result, the regret \(R^\pi(x, T; \lambda)\)
should be small.

To estimate \(p^D\), the algorithm dedicates an initial portion \([0, \tau]\) of the total
selling interval \([0, T]\) to an exploration of the price domain. On this initial interval,
the algorithm experiments with \(\kappa\) prices where each is kept fixed for \(\tau/\kappa\) units of
time. This structure leads to three main sources of error in the search for \(p^D\).
First, during the learning phase one incurs an exploration bias since the prices being tested there are not close to $p^D$ (or close to the optimal fixed price for that matter). This incurs losses of order $\tau$. Second, experimenting with only a finite number of prices $\kappa$ in the search for $p^D$ results in a deterministic error of order $1/\kappa$ in Step 3. Finally, only "noisy" demand observations are available at each of the $\kappa$ price points, and the longer a price is held fixed, the more accurate the estimate of the mean demand rate at that price. This introduces a stochastic error of order $(\tau/\kappa)^{-1/2}$ stemming from the nature of the Poisson process. The crux of the matter is to balance these three error sources by a suitable choice of the tuning parameters $\tau$ and $\kappa$.

3.4.2 Model uncertainty: the class of demand functions

The nonparametric class of functions we consider consists of regular demand function (satisfying the standard conditions laid out in Section 3.3) which in addition satisfy the following.

Assumption 5 For some finite positive constants $M, K, \bar{K}, m$, with $K \leq \bar{K}$:

(i.) Boundedness: $|\lambda(p)| \leq M$ for all $p \in [\underline{p}, \bar{p}]$.

(ii.) Lipschitz continuity: $|\lambda(p) - \lambda(p')| \leq K|p - p'|$ for all $p, p' \in [\underline{p}, \bar{p}]$ and $|\gamma(l) - \gamma(l')| \leq K^{-1}|l - l'|$ for all $l, l' \in [\lambda(\underline{p}), \lambda(\bar{p})]$.

(iii.) Minimum revenue rate: $\max\{p\lambda(p) : p \in [\underline{p}, \bar{p}]\} \geq m$.

Let $\mathcal{L} := \mathcal{L}(M, K, \bar{K}, m)$ denote this class of demand functions. Assumption 1(i.) and 1(ii.) are quite benign, only requiring minimal smoothness of the demand func-
tion; (ii.) also ensures that when \( \lambda(\cdot) \) is positive, it does not have “flat regions.”

Assumption 5(iii.) states that a minimal revenue rate exists, hence avoiding trivialities.

Note that assumptions (i.)-(iii.) above hold for many models of the demand function used in the revenue management and economics literature (e.g., linear, exponential and iso-elastic/Pareto with parameters lying in a compact set; cf. Talluri and van Ryzin (2005, §7) for further examples).

### 3.4.3 Performance analysis

Since minimax regret is hardly a tractable quantity, we introduce in this section an asymptotic regime characterized by a “high volume of sales,” which will be used to analyze the performance of Algorithm 1. We consider a regime in which both the size of the initial inventory as well as potential demand grow proportionally large. In particular for a market of “size” \( n \), where \( n \) is a positive integer, the initial inventory and the demand function are now assumed to be given by

\[
x_n = nx \quad \text{and} \quad \lambda_n(\cdot) = n\lambda(\cdot).
\]  

Thus, the index \( n \) determines the order of magnitude of both inventory and rate of demand. We will denote by \( \mathcal{P}_n \) the set of admissible policies for a market of scale \( n \), and the expected revenues under a policy \( \pi_n \in \mathcal{P}_n \) will be denoted \( J^n_n(x, T; \lambda) \). With some abuse of notation, we will occasionally use \( \pi \) to denote a sequence of policies \( \{\pi_n, n = 1, 2, \ldots\} \) as well as any element of that sequence, omitting the subscript “\( n \)” to avoid cluttering the notation. For each \( n = 1, 2, \ldots \), we denote by
\[ J_n^D(x, T|\lambda) \] the value of the deterministic relaxation given in (3.3.5) with the scaling given in (3.4.3); it is straightforward to verify that \( J_n^D(x, T|\lambda) = n J^D(x, T|\lambda) \).

Finally let the regret be denoted as \( R_n(x, T; \lambda) := 1 - J_n^x(x, T; \lambda)/J_n^D(x, T|\lambda) \). The following definition characterizes admissible policies that have “good” asymptotic properties.

**Definition 1 (Asymptotic optimality)** A sequence of admissible policies \( \pi_n \in \mathcal{P}_n \) is said to be asymptotically optimal if

\[
\sup_{\lambda \in \mathcal{L}} R_n(x, T; \lambda) \to 0 \quad \text{as } n \to \infty. \tag{3.4.4}
\]

In other words, asymptotically optimal policies achieve the full information upper bound on revenues as \( n \to \infty \), uniformly over the class of admissible demand functions.

For the purpose of asymptotic analysis we use the following notation: for real valued positive sequences \( \{a_n\} \) and \( \{b_n\} \) we write \( a_n = O(b_n) \) if \( a_n/b_n \) is bounded from above by a constant, and if \( a_n/b_n \) is also bounded from below then we write \( a_n \asymp b_n \). We now analyze the performance of policies associated with Algorithm 1 in the asymptotic regime described above.

**Proposition 1** Let Assumption 5 hold. Set \( \tau_n \asymp n^{-1/4} \), \( \kappa_n \asymp n^{1/4} \) and let \( \pi_n := \pi(\tau_n, \kappa_n) \) be given by Algorithm 1. Then, the sequence \( \{\pi_n\} \) is asymptotically optimal, and for all \( n \geq 1 \)

\[
\sup_{\lambda \in \mathcal{L}} R_n(x, T; \lambda) \leq \frac{C (\log n)^{1/2}}{n^{1/4}}, \tag{3.4.5}
\]

for some finite positive constant \( C \).
The constant $C$ above depends only on the parameters characterizing the class $\mathcal{L}$, the initial inventory $x$, and the time horizon $T$. The exact dependence is somewhat complex and is omitted, however, we note that it is fully consistent with basic intuition: as one expands the class $\mathcal{L}$ by suitably increasing or decreasing the value of the parameters in Assumption 1, the magnitude of $C$ grows, and vice versa. (This can also be inferred by carefully inspecting the proof of the proposition.)

We next present a lower bound on the minimax regret which establishes a fundamental limit on the performance of any admissible pricing policy. Roughly speaking, the main idea behind this result is to construct a “worst case” demand function such that the regret is large for any policy.

**Proposition 2** Let Assumption 5 hold with $M, K$ satisfying $M \geq \max\{2Kp, Kp + x/T\}$. Then there exists a finite positive constant $C'$ such that for any sequence of admissible policies $\{\pi_n\}$ and for all $n \geq 1$

$$\sup_{\lambda \in \mathcal{L}} R_n^\pi(x, T; \lambda) \geq \frac{C'}{n^{1/2}}.$$  \hspace{1cm} (3.4.6)

Combining Propositions 1 and 2, one can characterize the magnitude of the minimax regret as follows:

$$\frac{C'}{n^{1/2}} \leq \inf_{\pi \in \mathcal{P}_n} \sup_{\lambda \in \mathcal{L}} R_n^\pi(x, T; \lambda) \leq \frac{C(\log n)^{1/2}}{n^{1/4}},$$  \hspace{1cm} (3.4.7)

and hence the performance of Algorithm 1 is “not far” from being minimax optimal (i.e., achieving the lower bound). A question that remains open is whether one can close this gap by further refining the algorithm. We revisit this point in Sections 3.5 and 3.7.
3.5 Main Results: The Parametric Case

In this section we assume the demand function is known to have a parametric form. Our goal is to develop pricing policies that exploit this information and that will work well for a large class of admissible parametric demand functions. One of the main questions of interest here is whether these policies achieve a smaller regret relative to the nonparametric case, and if there exist policies whose performance cannot be improved upon.

**Preliminaries.** Let \( k \) denote a positive integer and \( \mathcal{L}(\Theta) = \{ \lambda(\cdot; \theta) : \theta \in \Theta \} \) be a parametric family of demand functions, where \( \Theta \subseteq \mathbb{R}^k \) is assumed to be a convex compact set, \( \theta \in \Theta \) is a parameter vector, and \( \lambda : \mathbb{R}_+ \times \Theta \to \mathbb{R}_+ \). We consider all parametric families that are subsets of the class of admissible regular demand functions defined in Section 3.4, i.e., \( \mathcal{L}(\Theta) \subseteq \mathcal{L} := \mathcal{L}(M, K, \overline{K}, m) \).

In the current setting, we will assume that the decision maker knows that \( \lambda \in \mathcal{L}(\Theta) \), i.e., s/he knows the parametric structure of the demand function, but does not know the value of the parameter vector \( \theta \). We continue to denote by \( \mathcal{P} \) the set of admissible pricing policies, i.e., policies that satisfy (3.3.2)-(3.3.3). For any policy \( \pi \in \mathcal{P} \), let \( J^\pi(x, T; \theta) \) denote the expected revenues under \( \pi \), and let \( J^D(x, T|\theta) \) denote the value of the deterministic relaxation (3.3.5) when the value of the unknown parameter vector is revealed to the decision maker prior to the start of the selling season. Let \( \mathcal{R}^\pi(x, T; \theta) \) denote the regret under a policy \( \pi \),
namely

\[ R^*(x, T; \theta) = 1 - \frac{J^*(x, T; \theta)}{J^D(x, T|\theta)}. \]

(3.5.1)

### 3.5.1 The proposed method

We consider a simple modification of the approach taken in Algorithm 1 that now exploits the assumed parametric structure of the demand function.

---

**Algorithm 2: \( \pi(\tau) \)**

**Step 1. Initialization:**

(a) Set the learning interval to be \([0, \tau]\). Put \( \Delta = \tau/k \), where \( k \) is the dimension of the parameter space \( \Theta \).

(b) Choose a set of \( k \) prices, \( \{p_i, i = 1, \ldots, k\} \).

**Step 2. Learning/experimentation:**

(a) On the interval \([0, \tau]\) apply \( p_i \) from \( t_{i-1} = (i - 1)\Delta \) to \( t_i = i\Delta \), \( i = 1, \ldots, k \), as long as the inventory is positive. If no more units are in stock, apply \( p_\infty \) up until time \( T \) and STOP.

(b) Compute

\[ \hat{d}_i = \frac{\text{total demand over } [t_{i-1}, t_i]}{\Delta}, \quad i = 1, \ldots, k. \]

(c) Let \( \hat{\theta} \) be a solution of \( \{\lambda(p_i; \theta) = \hat{d}_i, i = 1, \ldots, k\} \).
Step 3. Optimization:

\[ p^u(\hat{\theta}) = \arg\max\{p\lambda(p;\hat{\theta}) : p \in [p, \bar{p}]\}, \]

\[ p^c(\hat{\theta}) = \arg\min\{|\lambda(p;\hat{\theta}) - x/T| : p \in [p, \bar{p}]\}, \]

set \( \hat{p} = \max\{p^u(\hat{\theta}), p^c(\hat{\theta})\}. \)

Step 4. Pricing:

On the interval \((\tau, T]\) apply \(\hat{p}\) as long as inventory is positive, then apply \(p_{\infty}\) for the remaining time.

Note that Step 1(b) requires one to define the prices \(\{p_1, \ldots, p_k\}\) and Step 2(c) implicitly assumes that the system of equations admits a solution. (We define the prices and state this assumption more formally when analyzing the performance of Algorithm 2.) The intuition behind this algorithm is similar to the one that underlies the construction of Algorithm 1, the only difference being that the parametric structure allows one to infer accurate information about the demand function using only a "small" number of "test" prices \(k\). In particular, recalling the discussion following Algorithm 1, a key difference is that the deterministic error source associated with price granularity does not affect the performance of Algorithm 2.

3.5.2 The parametric class of demand functions

First note that the inclusion \(\mathcal{L}(\Theta) \subseteq \mathcal{L}\) implies that \(\lambda(\cdot; \theta)\) has an inverse for all \(\theta \in \Theta\); this inverse will be denoted \(\gamma(\cdot; \theta)\). (Note also that conditions (i.)-(iii.)
in Assumption 5 hold.) For any parameter vector \( \theta \in \Theta \) we denote the revenue function by \( r(l; \theta) := \gamma(l; \theta)l \). Let \( \{P_\theta : \theta \in \Theta \} \) be the family of demand distributions corresponding to Poisson processes with controlled intensities \( \lambda(\cdot; \theta), \theta \in \Theta \). We denote by \( f_{\lambda(p; \theta)}(\cdot) \) the probability mass function of a Poisson random variable with intensity \( \lambda(p; \theta) \).

The following technical conditions articulate standard regularity assumptions in the context of Maximum Likelihood estimation (cf. Borovkov (1998)), and are used to define the admissible class of parametric demand functions.

Assumption 6

(i.) There exists a vector of distinct prices \( \vec{p} = (p_1, ..., p_k) \in [\underline{p}, \overline{p}] \) such that:

a) For some \( l_0 > 0 \), \( \min_{1 \leq i \leq k} \inf_{\theta \in \Theta} \lambda(p_i; \theta) > l_0 \).

b) For any vector \( \vec{d} = (d_1, ..., d_k) \), the system of equations \( \{\lambda(p_i; \theta) = d_i, i = 1, ..., k\} \) has a unique solution in \( \theta \). Let \( g(\vec{p}, \vec{d}) \) denote this solution.

We assume in addition that \( g(\vec{p}, \cdot) \) is Lipschitz continuous with constant \( \alpha > 0 \).

c) For \( i = 1, ..., k \), \( \sqrt{\lambda(p_i; \theta)} \) is differentiable on \( \Theta \).

(ii.) For some \( K_2 > 0 \), \( |\lambda(p; \theta) - \lambda(p; \theta')| \leq K_2 \|\theta - \theta'\|_\infty \) for all \( p \in [\underline{p}, \overline{p}] \) and \( \theta, \theta' \in \Theta \).

Condition (i.) ensures that the parametric model is identifiable based on a sufficient set of observations. Condition (ii.) is a mild regularity assumption on the
parametric class, controlling for changes in the demand curve as parameters vary.

We provide below an example of a parametric class that satisfies Assumption 2.

**Example 1 (Linear demand function)** Let \( \lambda(p; \theta) = \theta_1 - \theta_2 p \), and set \([p_1, p] = [1, 2] \). Let \( \mathcal{L}(\theta) = \{ \lambda(\cdot; \theta) : \theta \in [10, 20] \times [1, 4] \} \). If one sets \((p_1, p_2) = (1, 2)\), it is straightforward to verify conditions (i.a), (i.c) and that (ii.) is satisfied with \( \bar{K}_2 = 1 \). It is also easy to see that the unique solution associated with condition (i.b) is given by

\[
g(p_1, p_2, d_1, d_2) = \left( d_1 + \frac{p_1}{p_2 - p_1} [d_1 - d_2], \frac{1}{p_2 - p_1} [d_1 - d_2] \right)
\]

and \(g(p_1, p_2, d_1, d_2)\) is clearly Lipschitz continuous with respect to \((d_1, d_2)\).

### 3.5.3 Performance analysis

Suppose that the initial inventory level and the demand function are scaled according to (3.4.3), and denote the regret by \( R_n^\pi(x, T; \theta) = 1 - J_n^\pi(x, T; \theta)/J_n^D(x, T|\theta) \).

**Proposition 3** Let Assumptions 5 and 6 hold and let \( \{p_1, \ldots, p_k\} \) be as in Assumption 6(i.). Set \( \tau_n \approx n^{-1/3} \) and let \( \pi_n := \pi(\tau_n) \) be defined by Algorithm 2. Then the sequence of policies \( \{\pi_n\} \) is asymptotically optimal and satisfies

\[
\sup_{\theta \in \Theta} R_n^\pi(x, T; \theta) \leq \frac{C(\log n)^{1/2}}{n^{1/3}}, \quad (3.5.2)
\]

for all \( n \geq 1 \) and some finite positive constant \( C \).

Contrasting the above with Proposition 1, we observe that the parametric structure of the demand function translates into improved performance bounds for
the learning and pricing algorithm that is designed with this knowledge in mind. In particular, the regret is now of order \( R_n = O((\log n)^{1/2} n^{-1/3}) \) as opposed to \( R_n = O((\log n)^{1/2} n^{-1/4}) \) in the nonparametric case. In addition, note that the upper bound above holds for all admissible parametric families. The improvement in terms of generated revenues, as quantified by the smaller magnitude of the regret, spells out the advantages of using a parametric approach versus a nonparametric one. The downside, namely, model misspecification and its consequences, is illustrated and discussed in Section 3.6.

At an intuitive level, classical estimation theory tells us that parameter uncertainty cannot be resolved faster than rate \( n^{-1/2} \) with \( n \) observations. This suggests that for the asymptotic regime we consider in this chapter, no admissible pricing policy would be able to achieve a convergence rate faster than \( n^{-1/2} \). This intuition is made rigorous in the next proposition.

**Proposition 4** Let Assumption 5 hold with \( M, \overline{K} \) satisfying \( M \geq \max\{2\overline{K} \overline{p}, \overline{K} \overline{p} + x/T\} \). Then, there exists a parametric family \( \mathcal{L}(\Theta) \subset \mathcal{L}(M, K, \overline{K}, m) \) satisfying Assumption 6 such that for some positive constant \( C' \) and for all admissible policies \( \{\pi_n\} \)

\[
\sup_{\theta \in \Theta} R_n^*(x, T; \theta) \geq \frac{C'}{n^{1/2}}, \quad (3.5.3)
\]

for all \( n \geq 1 \).

This result follows in a relatively straightforward manner from Proposition 2.
3.5.4 An optimal algorithm when a single parameter is unknown

Given the lower bound in (3.5.3), the natural question that arises is whether it is possible to close the remaining gap with respect to the upper bound in Proposition 3. We study this question in the context of a single unknown parameter, i.e., $k = 1$ and hence $\Theta \subseteq \mathbb{R}$. Consider the following $\ell$-step policy $\pi(\ell, \Delta^{(1)}, \ldots, \Delta^{(\ell)})$.

Algorithm 3: $\pi(\ell, \Delta^{(1)}, \ldots, \Delta^{(\ell)})$

Step 1. Initialization:

(a) Set the number of steps to be $\ell$ and define $\Delta^{(i)}$, $i = 1, \ldots, \ell$ so that

$$\Delta^{(1)} + \ldots + \Delta^{(\ell)} = T.$$

(b) Choose a price $\hat{p}_1 \in [\underline{p}, \bar{p}]$ as in Assumption 6(i).

Step 2. Learning/Optimization/Pricing:

Set $t_1 = 0$.

For $i = 1, \ldots, \ell$,

(a) Learning/Pricing:

i.) Apply $\hat{p}_i$ on the interval $[t_i, t_i + \Delta^{(i)}]$ as long as the inventory is positive. If no more units are in stock, apply $p_\infty$ up until time $T$ and STOP.
ii.) Compute
\[ \hat{d}_i = \frac{\text{Total demand over } [t_i, t_i + \Delta^{(i)}]}{\Delta^{(i)}}. \]

iii.) Set \( t_{i+1} = t_i + \Delta^{(i)} \).

iv.) Let \( \hat{\theta}^i \) be the unique solution of \( \lambda(p_i, \theta) = \hat{d}_i \).

(b) Optimization:
\[
\begin{align*}
p^\star(\hat{\theta}^i) &= \arg\max\{p\lambda(p; \hat{\theta}^i) : p \in [p, \bar{p}]\}, \\
p_c(\hat{\theta}^i) &= \arg\min\{|\lambda(p; \hat{\theta}^i) - x/t| : p \in [p, \bar{p}]\}, \\
\hat{p}_{i+1} &= \max\{p^\star(\hat{\theta}^i), p_c(\hat{\theta}^i)\}.
\end{align*}
\]

End For

The intuition underlying Algorithm 3 is as follows. When a single parameter is unknown, one can infer information about the parameter from observations of demand at a single price. Given this, the idea of Algorithm 3 is to price “close” to \( p^D \) after the first stage, but to continue learning. In particular, the estimate of \( p^D \) is improved from stage to stage using the demand observations from the previous stage. As a result, losses are mitigated by two effects: i.) after stage 1, the price is always close to \( p^D \); and ii.) the estimate of \( p^D \) becomes more precise. For what follows, we slightly strengthen Assumption 6(i.)

Assumption 7 \( \inf_{p \in [p, \bar{p}]} \inf_{\theta \in \Theta} \lambda(p; \theta) > l_0 \) and for any price \( p \in [p, \bar{p}] \) and any
\( d \geq 0, \) the equation \( \lambda(p; \cdot) = d \) has a unique solution. If \( g(p, d) \) denotes this solution, then \( g(p, \cdot) \) is Lipschitz continuous with constant \( \alpha > 0. \)

We now analyze the performance of policies associated with Algorithm 3. In particular suppose that the initial inventory level and the demand function are scaled according to (3.4.3), and define the sequence of tuning parameters \( \{\ell_n, \Delta_n^{(1)}, \ldots, \Delta_n^{(\ell_n)}\} \) as follows:

\[
\ell_n = (\log 2)^{-1} \log \log n \tag{3.5.4}
\]

\[
\Delta_n^{(m)} = \beta_n n^{(a_n/a_m)^{-1}}, \quad m = 1, \ldots, \ell_n, \tag{3.5.5}
\]

where \( a_m = 2^{m-1}/(2^m - 1) \) for \( m \geq 1 \), and \( \beta_n > 0 \) is a normalizing constant chosen so \( \Delta_n^{(1)} + \ldots + \Delta_n^{(\ell_n)} = T \). We then have the following result.

**Proposition 5** Let Assumptions 5,6 and 7 hold. Let \( \{\ell_n, \Delta_n^{(1)}, \ldots, \Delta_n^{(\ell_n)}\} \) be defined as in (3.5.4)-(3.5.5) and put \( \pi_n := \pi(\ell_n, \Delta_n^{(1)}, \ldots, \Delta_n^{(\ell_n)}) \), defined by Algorithm 3. Then the sequence of policies \( \{\pi_n\} \) is asymptotically optimal and satisfies

\[
\sup_{\theta \in \Theta} \mathcal{R}_n^\pi(x, T; \theta) = O\left(\frac{(\log \log n)(\log n)^{1/2}}{n^{1/2}}\right). \tag{3.5.6}
\]

Note that the regret of the sequence for policies introduced in Proposition 5 achieves the lower bound spelled out in Proposition 4 (up to logarithmic terms). In that sense, these policies cannot be improved upon. On the other hand, Algorithm 3 is restricted to the case where only one parameter is unknown, and exploits the fact that in this setting it is possible to learn the single unknown parameter by conducting price experiments in the neighborhood of the near-optimal price \( p^D \).
Designing policies that achieve the lower bound in Proposition 4 in the multi-parameter case remains an open question.

3.6 Numerical Results and Qualitative Insights

3.6.1 The "price" of uncertainty

We examine the performance of three policies developed in the previous sections:

i.) The nonparametric policy defined in Algorithm 1; ii.) The parametric policy defined by means of Algorithm 2 and designed for a finite number of unknown parameters; and iii.) The parametric policy defined in Algorithm 3 designed for cases with a single unknown parameter. (The tuning parameters are taken as in Propositions 1, 3 and 5, respectively.)

The performance of these policies are measured by the magnitude of the regret. Note that $R_n(x, T; \lambda) \approx C/n^\gamma$ implies that $\log R_n(x, T; \lambda)$ should be approximately linear in $\log n$ with slope $-\gamma$. In Figure 3.1, we depict $R_n(x, T; \lambda)$ as a function of $n$ in a log-log plot for large values of $n$, and compute the best (least squares) linear fit for these values. The results depicted are based on running $10^3$ independent simulation replications from which the performance indicators were derived by averaging. The standard error for $R_n(x, T; \lambda)$ was below 0.05% in all cases.

Figure 3.1(a) summarizes results for an underlying exponential demand model $\lambda(p) = \theta \exp(-.5p)$, where $\theta = 10 \exp(1)$, and Figure 3.1(b) presents results for an underlying linear demand model $\lambda(p) = 30 - \theta p$ where $\theta = 3$. The nonparametric algorithm does not make any assumptions with regard to the structure of the
demand function (beyond those spelled out in Section 4), while the parametric algorithms are assumed to know the parametric structure but the true value of $\theta$ is not revealed to them. In both cases, the initial normalized inventory level was $x = 20$, the selling horizon was $T = 1$ and the feasible price set was $[p, \bar{p}] = [0.1, 10]$.

![Performance of pricing policies as a function of the market size ($n$). Stars show the performance of the nonparametric policy defined in Algorithm 1; dots represent the performance of the parametric policy defined Algorithm in 2; and squares depict the performance of the parametric policy defined in Algorithm 3. The dashed lines represent the best linear fit to each set of points; in panel (a), the demand function is exponential and in panel (b) linear.](image_url)

**Discussion.** The slopes of the best linear fit in Figures 3.1(a) and 3.1(b) are very close to $\gamma = -1/4$, $\gamma = -1/3$ and $\gamma = -1/2$, predicted by the upper bounds in (3.4.5), (3.5.2) and (3.5.6), respectively. These results provide a "picture proof" of the Propositions 1, 3 and 5. As is evident, the less structure is assumed a priori, the higher the profit loss relative to the full information benchmark: informally
speaking, this is the "price" paid due to increasing uncertainty with regard to the demand model.

An additional fundamental difference between the various policies concerns the degree to which the price domain is explored. The nonparametric policy (Algorithm 1) essentially needs to explore the "entire" price domain; the general parametric policy (Algorithm 2) needs to test a number of prices equal to the number of unknown parameters; and the parametric policy designed for the case of a single unknown parameter (Algorithm 3) explores only prices in the vicinity of the price $p^D$. In other words, as uncertainty "decreases," the price exploration region shrinks as well.

3.6.2 The "price" of hedging misspecification risk

We have seen in the previous section that more refined information regarding the demand model yields higher revenues. Thus, it becomes tempting to postulate parametric structure, which in turn may be incorrect relative to the true underlying demand model. This misspecification risk can be eliminated via nonparametric pricing policies, but at the price of settling for more modest performance revenue-wise. We now provide an illustration of this trade-off.

We fix the time horizon to be $T = 1$ and the set of feasible prices to be $[p, \bar{p}] = [0.1, 10]$. Figure 3.2 depicts the regret $\mathcal{R}_n(x, T; \lambda)$ for two underlying demand models, various values of $n$, and three policies. The first policy assumes an exponential parametric structure for the demand function, $\lambda(p; \theta) = \theta_1 \exp(-\theta_2 p)$ with $\theta = (\theta_1, \theta_2)$, and applies Algorithm 2 with $\tau_n = n^{-1/3}$. The second policy
assumes a linear parametric structure \( \lambda(p; \theta) = (\theta_1 - \theta_2 p)^+ \) and applies Algorithm 2 with \( \tau_n = n^{-1/3} \). Finally, the third policy is the one given by the nonparametric method described in Algorithm 1 (with tuning parameters as in Proposition 1). The point of this comparison is to illustrate that the parametric algorithm outperforms its nonparametric counterpart when the assumed parametric model is consistent with the true demand function, otherwise the parametric algorithm leads to a regret that does not converge to zero due to a model misspecification error.

In Figures 3.2(a),(b) the underlying demand model is given by \( \lambda(p) = a \exp(-ap) \) with \( a = 10 \exp(1) \) and \( a = 1 \), and hence the policy that assumes the exponential structure (squares) is well specified, while the policy that assumes a linear demand function (crosses) suffers from model misspecification.

In Figures 3.2(c),(d) the underlying demand model is given by \( \lambda(p) = (a - ap)^+ \) with \( a = 30 \) and \( a = 3 \). Now the policy that assumes the exponential structure (squares) corresponds to a misspecified case, while the policy that assumes a linear demand function (crosses) corresponds to a well specified case.

In Figures 3.2(a),(c) the normalized inventory is \( x = 8 \) and in Figures 3.2(b),(d) it is taken to be \( x = 20 \). The results depicted in the figures are based on running \( 10^3 \) independent simulation replications from which the performance indicators were derived by averaging. The standard error was below 0.9% in all cases.

**Discussion.** We focus on Figures 3.2(a),(b) as similar remarks apply to Figures 3.2(c),(d). First, note that the parametric algorithm is able to achieve close to 90%
Figure 3.2: Performance of the parametric (Algorithm 2) and nonparametric policies (Algorithm 1) as a function of the market size: the parametric algorithm that assumes a linear model (crosses) and exponential model (squares); and the nonparametric algorithm (stars). The underlying demand model is exponential in Figures (a) and (b) and linear in Figures (c) and (d). Arrows indicate if the approach used is nonparametric [np] or parametric [p], and whether the model is well/misspecified in the case of the parametric algorithm.

of the full information revenues when the parametric model is well specified, when the market size is \( n = 100 \) or more. In addition, the convergence of the regret for Algorithm 2 is faster for small \( n \) under the well specified parametric assumption (squares) in comparison to the nonparametric algorithm (stars). [We note that the non-monotonic performance of the nonparametric algorithm (stars) stems from the
fact that the price grid used by this algorithm changes with \( n \), and consequently the minimal distance of the sought fixed price to any price on the grid need not be monotonic with respect to \( n \).]

In contrast, if the model is misspecified, the regret \( R_n(x, T; \lambda) \) converges to a strictly positive value (crosses), as expected, and hence the performance of the parametric algorithm fails to asymptotically achieve the full information revenues. The takeaway message here is that a nonparametric approach eliminates the risk stemming from model misspecification but at a price of extracting lower revenues than its parametric counterpart. The regret bounds derived in earlier sections precisely spell out the magnitude of this “price” of misspecification.

### 3.7 Sequential Learning Strategies in the Non-parametric Setting

The performance guarantees associated with Algorithm 1 rely heavily on how well one is able to estimate the two prices \( p^u \) and \( p^c \) arising from the solution of the deterministic relaxation (3.3.5). In Algorithm 1, this estimation is executed in a single stage over the interval \([0, \tau]\). A natural extension would be to consider policies in which the learning phase searches for the optimal price in a sequential and dynamic manner.

Recall the definitions of \( p^u \) and \( p^c \), the unconstrained and constrained solutions of the deterministic relaxation: \( p^u = \arg\max_{p \in [\underline{p}, \overline{p}]} \{ r(\lambda(p)) \} \), \( p^c = \arg\min_{p \in [\underline{p}, \overline{p}]} |\lambda(p) - x/T| \). Suppose for simplicity that the two critical prices belong to the interior of the feasible price set \([\underline{p}, \overline{p}]\). Then \( p^u \) is the unique maximizer of the revenue function
and $p^c$ is the $x/T$-crossing of the demand function, which is assumed to be decreasing. Using this interpretation, one could use stochastic approximation schemes to search for the two prices rather than via the estimation of the demand function over the whole price domain (as in Algorithm 1). We detail below how such schemes could be designed.

**Background on stochastic approximations.** Suppose one wants to estimate the maximizer $z^*$ of a real valued function $M(z)$ on a domain $[z, \bar{z}]$, and that the available observations are in the form $H(z) = M(z) + \epsilon(z)$, where $\epsilon(z)$ is a stochastic noise term with $\mathbb{E}[\epsilon(z)] = 0$ and $\mathbb{E}[\epsilon^2(z)] < \infty$ for all $z \in [z, \bar{z}]$. The basic Kiefer-Wolfowitz (KW) scheme estimates $z^*$ using the recursion

$$z_{k+1} = z_k + a_k \frac{y_{2k} - y_{2k-1}}{c_k}, \quad k = 1, 2, \ldots, (3.7.1)$$

where $y_{2k} = H(z_k + c_k)$ and $y_{2k-1} = H(z_k - c_k)$, and $\{a_k\}$ and $\{c_k\}$ are predefined deterministic real-valued positive sequences. Under suitable technical conditions on the function $M(\cdot)$ and the sequences $\{a_k\}$ and $\{c_k\}$, Kiefer and Wolfowitz (1952) showed that the sequence $z_k \rightarrow z^*$ in probability as $k \rightarrow \infty$. Similar ideas lead to a sequence $\{z_k\}$ that converges to an $\alpha$-crossing $\{z : M(z) = \alpha\}$ via the Robbins-Monro (RM) procedure (see Robbins and Monro (1951)), viz.

$$z_{k+1} = z_k + a_k (\alpha - y_k), \quad (3.7.2)$$

where $y_k = H(z_k)$. For further details and other variants the reader is referred to Benveniste, Métivier, and Priouret (1990).

**A stochastic approximation-based algorithm.** A straightforward way
to blend adaptive search procedures into Algorithm 1 is given below. Here \( h(\cdot) \)
denotes the projection operator \( h(y) := \min\{\max\{y, p\}, \bar{p}\} \).

Algorithm 4 : \( \pi^{SA}(\tau, \Delta, \{a_k\}, \{c_k\}) \)

Step 1. Initialization:

Set the learning interval to be \([0, \tau]\), and set \( \Delta \) to be the holding time for
each price during that period.

Step 2. Learning/experimentation:

(a) KW-type scheme:

Set \( \kappa_1 = \left\lfloor \frac{\tau}{4\Delta} \right\rfloor \), \( t_1 = 0 \), \( z_1 = p \) and \( t_i = t_1 + (i - 1)\Delta, \ i = 2, \ldots \)

For \( k = 1, \ldots, \kappa_1 \)

i.) Apply \( p_k = \min\{z_k + c_k, \bar{p}\} \) on \([t_{2k-2}, t_{2k-1})\) and compute

\[ y_{2k} = \left( \frac{\text{total demand over } [t_{2k-2}, t_{2k-1})}{\Delta} \right) \]

ii.) Apply \( p_k = \max\{z_k - c_k, p\} \) on \([t_{2k-1}, t_{2k})\) and compute

\[ y_{2k-1} = \left( \frac{\text{total demand over } [t_{2k-1}, t_{2k})}{\Delta} \right) \]

iii.) Update: \( z_{k+1} = h\left(z_k + \left(\frac{a_k}{c_k}\right)(y_{2k} - y_{2k-1})\right) \)

end For

Set \( \bar{p} = z_{\kappa_1+1} \)
(b) RM-type scheme:

Set $\kappa_2 = \lceil \tau/(2\Delta) \rceil$, $t_1 = 2\kappa_1 \Delta$, $z_1 = p$ and $t_i = t_1 + (i-1)\Delta$, $i = 2, \ldots$

For $k = 1, \ldots, \kappa_2$

i.) Apply $p_k = \min\{z_k, p\}$ on $[t_k, t_{k+1})$ and compute

$$y_k = \left(\text{total demand over } [t_k, t_{k+1})\right)/\Delta$$

ii.) Update: $z_{k+1} = h(z_k + a_k(x/T - y_k))$

end For

Set $\hat{p}^c = z_{\kappa_2+1}$

Step 3. Optimization: Set $\hat{p} = \max\{\hat{p}^u, \hat{p}^c\}$.

Step 4. Pricing: Set $\tau' = (2\kappa_1 + \kappa_2)\Delta$. On the interval $(\tau', T]$ apply $\hat{p}$ as long as

inventory is positive, then apply $p_\infty$ for the remaining time.

Comments on the structure of the algorithm. The algorithm exploits the structure of the single product problem, adaptively estimating each of the two prices $p^u$ and $p^c$ based on its characterization as either a point of maximum or a level crossing; the former is executed in Step 2(a) and the latter in the step 2(b). This also requires two consecutive price testing phases to be carried out. Note
that in Step 2 above, the number of iterations is different for the KW and RM procedures. This is due to the fact that a KW-based procedure uses two prices at each iteration in order to estimate an improvement direction. (For brevity, in Step 2 it was left implicit that as soon as inventory is depleted, the shut-off price $p_\infty$ is applied.)

**Numerical results.** We present a comparison between the performance of the policy defined by Algorithm 1 with tuning parameters as in Proposition 1, and the policy given by Algorithm 4 that uses stochastic approximations. We use $\pi_1^{SA}$ to denote the policy which uses Algorithm 4 taking $\tau_n = n^{-1/4}$ and $\Delta_n = n^{-1}$; this corresponds to the KW-scheme using $\left\lceil (1/4)n^{3/4} \right\rceil$ steps. Similarly, we let $\pi_2^{SA}$ denote the policy which corresponds to taking $\tau_n = n^{-1/4}$ and $\Delta_n = n^{-1/2}$. (The choice of $\tau_n$ is driven by our previous asymptotic considerations and is meant only for illustrative purposes.) The key difference between $\pi_1^{SA}$ and $\pi_2^{SA}$ is that in the latter there are fewer test prices being used, and the observations for a given price are less noisy since each price is held for a longer time interval.

Table 3.1 presents results for the three policies discussed above and for two normalized initial inventory levels ($x$); if the problem size is $n$ and the total inventory is $y$ then $x = y/n$. For each policy we record the regret $R_n^*(x, T; \lambda)$, and the number of prices $\kappa$ that are used during the interval $[0, T]$. To generate the results in Table 3.1 we take the underlying demand function to be linear, $\lambda(p) = (10 - 2p)^+$ and the set of feasible prices is taken to be $[p, \bar{p}] = [0.1, 4.5]$. In the first case where $x = 3$, and hence $p^c = 3.5$ and $p^u = 2.5$, the constrained price $p^c$ is optimal in
the deterministic relaxation (3.3.5). In the second case, there is a larger initial normalized inventory $x = 8$ and the prices of interest are given by $p^c = 1$ and $p^u = 2.5$, hence the unconstrained price is optimal in the deterministic relaxation.

<table>
<thead>
<tr>
<th>Problem “size”</th>
<th>$n = 10^2$</th>
<th>$n = 10^3$</th>
<th>$n = 10^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{R}_n(x, T; \lambda)$</td>
<td>$\kappa$</td>
<td>$\mathcal{R}_n(x, T; \lambda)$</td>
<td>$\kappa$</td>
</tr>
<tr>
<td>$x = 3$, $\pi$</td>
<td>.44</td>
<td>5</td>
<td>.19</td>
</tr>
<tr>
<td>$\pi_{SA}^1$</td>
<td>.25</td>
<td>33</td>
<td>.12</td>
</tr>
<tr>
<td>$\pi_{SA}^2$</td>
<td>.27</td>
<td>12</td>
<td>.12</td>
</tr>
<tr>
<td>$x = 8$, $\pi$</td>
<td>.86</td>
<td>5</td>
<td>.08</td>
</tr>
<tr>
<td>$\pi_{SA}^1$</td>
<td>.28</td>
<td>33</td>
<td>.17</td>
</tr>
<tr>
<td>$\pi_{SA}^2$</td>
<td>.22</td>
<td>12</td>
<td>.08</td>
</tr>
</tbody>
</table>

Table 3.1: Performance of stochastic approximation based schemes. $\mathcal{R}_n(x, T; \lambda)$ represents the regret associated with a given policy, and $x$ is the normalized inventory level. Here $\pi$ is the basic nonparametric pricing policy defined by Algorithm 1, $\pi_{SA}^1$ and $\pi_{SA}^2$ correspond to two implementations of Algorithm 4, and $\kappa$ = number of prices used by a given policy.

The results depicted in the table are based on running 500 independent simulation replications from which the performance indicators were derived by averaging. The standard error for $\mathcal{R}_n(x, T; \lambda)$ was below 0.8% in all cases. (We note that numerical experiments with other simple demand models give rise to similar results and therefore, for space considerations, we only report on the linear demand example.)

**Discussion.** A quick inspection of the results in Table 3.1 leads to the following observations. First, all three policies are seen to have comparable performance achieving a regret which is smaller than 20% for problems where the inventory is of the order of a thousand; this translates to excess of 80% of the optimal full information revenues. It is also interesting to contrast the structure of the two adaptive policies $\pi_{SA}^1$ and $\pi_{SA}^2$. We observe that the former uses significantly more
prices than the latter (180 prices for a problem with \( n = 10^3 \), compared with
18 for \( \pi_2^{SA} \)) and hence is not very practical to implement. Considering that the
performance of \( \pi_1^{SA} \) and \( \pi_2^{SA} \) is comparable, it seems that choosing the length \( \Delta \)
so that observations are less noisy is potentially preferable, even though less price
changes are allowed. The nonparametric policy \( \pi \) uses less than half of the price
changes that \( \pi_2^{SA} \) utilizes, and is overall the simplest to implement.

A detailed performance analysis of Algorithm 4 is beyond the scope of the
chapter. However, based on convergence rate results associated with stochastic
approximation schemes (see Polyak and Tsybakov (1990) and Polyak and Juditsky
(1990)), we expect that the policies associated with Algorithm 4 would achieve at
best a regret of order \( \mathcal{R}_n \approx n^{-1/3} \).

In Algorithm 4, we use a KW scheme that seeks an estimate of \( p^u \) (the un-
constrained maximizer of \( r(\cdot) \)) and an RM-based scheme that seeks an estimate
of \( p^c \) (the inventory-constrained "run out" price). Only after these estimates are
obtained, by running each stochastic approximation procedure separately, can one
attempt to deduce which of the two prices is "near-optimal" for the given problem.
An interesting direction is to investigate methods for finding an estimate of the
optimal price of the deterministic relaxation using a single combined procedure.
Stochastic approximations for constrained problems exist (see, e.g., Kushner and
Sanvicente (1975)), however, to the best of our knowledge there are still open
questions with regard to their performance and rates of convergence.
Chapter 4

Extensions to the Multi-product Case

4.1 Introduction

4.1.1 Background and overview of the main contributions

Among the first papers to propose a general mathematical model for network revenue management problem was that of Gallego and van Ryzin (1997). In their model they assume realized demand is given by a (multivariate) Poisson process whose instantaneous rate represents mean demand for each product type, and is controlled by a vector of prices chosen by the decision maker. Given an initial inventory of primitive resources used to construct the finished products, the objective is to maximize the total expected revenues over the course of a selling season. For this, the optimal dynamic pricing policy can be obtained, at least in principle, by exploiting Markovian structure and solving the associated Bellman equation. Roughly speaking, the resulting policy seeks to adjust prices at each point in time based on current inventory levels so as to maximize future expected profits.
The dynamic programming logic articulated above, and variants thereof, form the analytical backbone of most revenue management studies to date (cf. Talluri and van Ryzin (2005)). The vast majority of these studies are predicated on the assumption of "full information," namely, that the demand function, i.e., the functional relationship that determines how price affects mean demand rate, is known to the decision maker at the start of the selling season. The only remaining source of uncertainty is the randomness of realized demand. Needless to say, this type of stipulation is invalid in most practical settings where one does not possess accurate information, if at all, on the nature of the demand function.

The present paper is, to the best of our knowledge, among the first to propose and analyze a general approach for solving dynamic pricing problems on a network, without any significant prior information on the demand model; see further discussion in the next section that reviews relevant work. The main building blocks that are used here are quite distinct from almost all antecedent literature on the topic, in particular, they are not guided by dynamic programming principles. Rather we pursue a blend of ideas from nonparametric estimation and large scale system analysis to derive efficient pricing policies, that perform well by suitably balancing exploration-exploitation tradeoffs which are intrinsic to the problem at hand.

Overview of the main contributions and qualitative insights. The aim of our work is to extend the classical network revenue management setting of Gallego and van Ryzin (1997) to one where the demand function is unknown and little is assumed insofar as its properties (in particular, it need not admit
a parametric representation). For this reason we refer to the class of problems studied in this paper as “blind” network revenue management, to indicate the opaque nature of information available to the decision maker.

The problem described above falls into a broader setting whose focal point is dynamic optimization under model uncertainty. A key challenge here, in particular in the context of the problem of blind network revenue management, is the curse of dimensionality. To elucidate this point, it is beneficial (and mathematically appealing) to first consider the case in which the action space in our control problem is finite. That is, we will first consider the case where the set of feasible prices is finite and fixed in advance. Such instances arise, for example, in settings where prices need to be defined as given percentages of a “base” price (see also, e.g., Talluri and van Ryzin (2005, §5.2.1.3)).

In this context, we develop a simple linear programming-based policy, summarized in Algorithm 5, which uses an initial learning phase to determine “how long” to hold each price so as to approximately maximize revenues. We prove that this policy is asymptotically optimal in the following sense: as the volume of sales grows large, the revenue it generates approaches the revenues extracted by the optimal dynamic pricing policy that knows the demand function a priori; see Theorem 4. Moreover the performance of the policy does not depend on the number of products being sold, i.e., it is dimension-independent. This is essentially seen to be a consequence of the finiteness of the set of feasible actions/prices.

We then consider the more complicated case in which the action set is un-
countable, namely, there is a continuum of feasible prices from which one needs
to construct the optimal pricing policy. The main problems here are: i.) how to
select a suitable and sufficiently rich subset of “test prices” for purposes of learn-
ing demand response; and ii.) how long to experiment with each price in order
to assure accurate estimates of demand. Since the unknown demand function is
multidimensional, the number of prices being tested needs to be large. On the
other hand, since only realized (“noisy”) demand is observable, accurate estimates
of the demand rate require each price to be applied for a sufficient amount of time.
But throughout this time, extracted revenues are most likely suboptimal. These
contradicting objectives need to be balanced for a policy to perform well.

We first develop a simple policy that tests a discrete subset of prices in the
exploration (demand learning) phase, and then selects the “best” price to be used
in the exploitation (pricing and revenue extraction) phase; see Algorithm 6. Our
analysis establishes that the policy is asymptotically optimal, but at the same time,
its performance degrades significantly with the number of products being sold as
a consequence of the curse of dimensionality; see Theorem 5.

We then propose a modification of this policy that uses the demand data ob-
tained in the price testing phase to construct a nonparametric estimate of the
entire demand function and revenue surface. This functional estimate is then fed
into a deterministic optimization problem which gives rise to the ultimate pricing
policy; see Algorithm 6’.

The policy described above exploits prior knowledge on the smoothness of the
demand function to guide both data collection (price testing), and the nonparametric curve fitting stages. Unlike most work in the full information setting, where smoothness is typically imposed as a purely technical condition, in our context smoothness plays a much more instrumental role: it communicates important information on the unknown demand function. Roughly speaking, the smoother the demand surface, the less one suffers from dimensionality effects; this is articulated in precise mathematical terms in Theorem 6 (see also Remark 1 following the theorem).

An important implication of the above is that with extra smoothness imposed on the unknown demand function, good performance can be achieved while requiring only a "moderate" number of prices to be tested in the learning phase. The intuition here is that a smooth function can be "pinned down" using a smaller number of test points, in contrast to a function that varies more sharply; Theorem 6 and Remark 2 quantify this point and put it on a rigorous foundation. This makes the approach proposed here appealing also from a practical perspective: in many settings, especially when prices are visible to all consumers, it is preferable to test as few prices as possible.

A final comment concerns the learning phase in our pricing strategies, which is closely related to the widespread industry practice of "price testing." A recent empirical study of 32 large U.S. retailers, finds that nearly 90% of them conduct such price experiments (see Gaur and Fisher (2005)), and the advent of the Internet and the Direct-to-Customer model have served to greatly facilitate such price test-
ing practices and their implementation (see, e.g., Williams and Partani (2006) for further discussion and examples). Given the central role of price testing practices, there is a growing need to better understand this approach and add to its rigorous foundations. (It is worth noting that the use of “testing” ideas is not limited to prices; see Fisher and Rajaram (2000) for a study involving merchandise testing.)

Our analysis attempts to shed some light on this issue by providing simple and intuitive guidelines for selecting both the number of prices that should be tested, as well as the overall fraction of the selling season that should be dedicated to experimentation.

4.1.2 Related literature

Almost all work we are aware of that incorporates model uncertainty into the dynamic pricing problem described above, has effectively been restricted to the one dimensional case (where there is only a single product being sold). The bulk of these studies focus on a parametric setting where the structure of the demand function is assumed to be known up to a finite number of unknown parameters. The method of choice in the analysis of such problems has been a Bayesian formulation of dynamic programming; see Lobo and Boyd (2003), Aviv and Pazgal (2005), Araman and Caldentey (2005), and Farias and Van Roy (2006), all of which restrict attention to one or two unknown parameters. Distinct from this stream of literature is the analysis of Chapter c:single that proposes a “frequentist” approach to the problem, using maximum likelihood to infer the unknown parameters, and policies that hinge on a separation of estimation and control.
For any parametric approach to work well, it is crucial that the structure assumed by the policy be consistent with that of the true underlying demand function. In other words, the postulated model needs to be well specified with respect to the actual mechanism that determines realized demand. To remove misspecification risk, one needs to step outside the boundaries of parametric modeling assumptions. For example, one can assume that the unknown demand function satisfies some mild nonparametric structural conditions (e.g., that it is monotone, bounded, differentiable, etc.). Very little work has been done to date in this direction. A few recent studies consider static settings, which do not involve dynamic decision making over time and tradeoffs between learning and pricing; see, e.g., Rusmevichientong et al. (2006) and Eren and Maglaras (2007) (see also van Ryzin and McGill (2000) and Ball and Queyranne (2006) in the context of capacity allocation problems). An exception is the work of Lim and Shanthikumar (2006) that formulates a robust max-min analogue of the dynamic pricing problem of Gallego and van Ryzin (1994); see also Lim et al. (2006) for an analysis of the multiproduct case. Their work is fairly conservative insofar as an adversary (nature) is allowed to alter the distribution of realized demand at each point in time to counter any chosen policy, and with the exception of exceedingly simple cases, the approach is not tractable and does not lead to prescriptive solutions.

In the present chapter, we focus exclusively on the setting where the demand function is not assumed to possess any parametric structure. The main issue in the current network setting, which does not arise at all in the single product problem,
is the so-called “curse of dimensionality,” further exacerbated by the lack of prior information on the demand function. This is one of the main focal points of the present paper.

The remainder of the chapter. The next section introduces the model and formulates the problem. Section 4.3 analyzes the blind network problem where the feasible price set is discrete and finite. Section 4.4 shifts focus to the general blind network case, and Section 4.5 analyzes ways to exploit further structural assumptions about the demand function to counter dimensionality effects. Section 4.6 contains some concluding remarks and discussion of the modeling assumptions. All proofs are collected in two appendices: Appendix C.1 contains the proofs of the main results and Appendix C.2 contains proofs of auxiliary lemmas.

4.2 Problem Formulation

The model. We consider a revenue management problem in which a firm sells \( d \) different products which are generated (assembled or produced) from \( \ell \) resources. Let \( A = [a_{ij}] \) denote the capacity consumption matrix, whose entries \( a_{ij} \geq 0 \), \( i = 1, \ldots, \ell \) and \( j = 1, \ldots, d \), denote the number of units of resource \( i \) required to generate product \( j \). It is assumed that the entries of \( A \) are integer valued and each column contains at least one non-zero entry. The selling horizon is denoted by \( T > 0 \), and after this time sales are discontinued and there is no salvage value for the remaining unsold products.

Demand for products at any time \( t \in [0, T] \) is given by a multivariate Poisson
process with intensity $\lambda_t = (\lambda_1^t, \ldots, \lambda_d^t)$ which measures the instantaneous demand rate (in units such as number of products requested per hour, say). This intensity is determined by the price vector at time $t$, $p(t) = (p^1(t), \ldots, p^d(t))$ through a demand function $\lambda : \mathcal{D}_p \rightarrow \mathbb{R}_+^d$, where $\mathcal{D}_p \subseteq \mathbb{R}_+^d$ denotes the set of feasible prices. Thus the instantaneous demand rate at time $t$ is given by $\lambda_t = \lambda(p(t))$, and the realized demand is a controlled Poisson process. More will be said on the demand function shortly.

Let $(p(t) : 0 \leq t \leq T)$ denote the price process which is assumed to have sample paths that are right continuous with left limits taking values in $\mathcal{D}_p$. Let $(N^1(\cdot), \ldots, N^d(\cdot))$ be a vector of mutually independent unit rate Poisson processes. The cumulative demand for product $j$ up until time $t$ is then given by $D^j(t) := \int_0^t \lambda^j(p(s))ds$. We say that $(p(t) : 0 \leq t \leq T)$ is non anticipating if the value of $p(t)$ at time $t \in [0, T]$ is only allowed to depend on past prices $\{p(s) : s \in [0, t]\}$ and demand values $\{(D^1(s), \ldots, D^d(s)) : s \in [0, t]\}$. (That is, the price process is adapted to the filtration generated by past values of the demand and price processes.)

**Information structure and the dynamic optimization problem.** We assume that the decision-maker does not know the true demand function and only knows that $\lambda$ belongs to the class $\mathcal{L} := \mathcal{L}(M, m, p_\infty)$, which for finite positive constants $M$, $m$ and a vector $p_\infty \in \mathcal{D}_p$ satisfies the following:

i.) Boundedness of demand: for all $\lambda \in \mathcal{L}$, $\|\lambda(p)\| \leq M$ for all $p \in \mathcal{D}_p$. 
ii.) Minimum revenue rate: for all $\lambda \in \mathcal{L}$, $\max\{p \cdot \lambda(p) : p \in \mathcal{D}_p\} \geq m$.

iii.) "Shut-off" price: for all $\lambda \in \mathcal{L}$, $\lambda(p_\infty) = 0$.

Here for two vectors $y, z \in \mathbb{R}^d$, $y \cdot z$ denotes the usual scalar product and $\|y\| := \max\{|y^i| : i = 1, \ldots, d\}$. It is worth noting that Assumptions i.) and ii.) are quite benign and hold for many demand models used in the revenue management literature such as linear, exponential and iso-elastic (Pareto), as long as the parameters are assumed to lie in a compact set; see, e.g., Talluri and van Ryzin (2005, §7) for further examples. The existence of a "shut-off" price in Assumption iii.) is not restrictive from a practical standpoint since in most applications there exists a finite price that yields zero demand. From a modeling perspective, this is merely a convenient way to allow for a sales denial.

While the decision maker possesses only limited information on the demand function, s/he is able to continuously observe realized demand at all time instants starting at time 0 and up until the end of the selling horizon $T$. We shall use $\pi$ to denote a pricing policy and its associated price process will be denoted $(p(t) : 0 \leq t \leq T)$. With some abuse of terminology, we will use the term policy to refer to the price process itself, as well as the algorithm that generates it interchangeably.

For $0 \leq t \leq T$ put

$$N^{j,\pi}(t) := N_j^j \left( \int_0^t \lambda_j(p(s)) ds \right), \quad \text{for } j = 1, \ldots, d,$$

(4.2.1)

where $N^{j,\pi}(t)$ denotes the cumulative demand, i.e., number of units requested of product $j$ up to time $t$ under the policy $\pi$. Let $N^\pi(t)$ denote the vector
Let \( x = (x_1, x_2, \ldots, x^\ell) \) denote the inventory level of each resource at the start of the selling season. We assume without loss of generality that \( x^i > 0, i = 1, \ldots, \ell \).

A policy \( \pi \) is said to be admissible if the induced price process is non-anticipating and satisfies

\[
\int_0^T A dN^\pi(s) \leq x \quad \text{a.s.,} \\
p(s) \in D_p, \quad 0 \leq s \leq T,
\]

where \( A \) is the capacity consumption matrix defined earlier and vector inequalities are assumed to hold componentwise. The term non-anticipating means that at any point in time, \( p(t) \) can only demand on past realized demand \( (N^\pi(s) : 0 \leq s < t) \) and prices \( (p(s) : 0 \leq s < t) \). It is important to note that while the decision maker does not know the demand function, knowledge of \( p^\infty \) guarantees that the constraint (4.2.2) can be met. We let \( \mathcal{P} \) denote the set of admissible policies, and the performance of a policy \( \pi \in \mathcal{P} \) is measured in terms of cumulative expected revenues,

\[
J^\pi(x, T; \lambda) := \mathbb{E} \left[ \int_0^T p(s) \cdot dN^\pi(s) \right].
\]

It is worth noting that the decision maker is not able to compute the expectation in (4.2.4) since the true demand function governing customer requests is not known a priori. This lends further meaning to the terminology “blind revenue management,” where one is attempting to optimize (4.2.4) in a blind manner.

**The full information benchmark and main objective.** When the de-
mand function \( \lambda \) is known prior to the start of the selling season, the dynamic optimization problem described above can, at least in theory, be solved; this will be referred to as the “full information” setting. This problem is precisely the one formulated in Gallego and van Ryzin (1997), who also characterize the optimal state-dependent pricing policy using dynamic programming. Suppose that we fix a demand function \( \lambda \in \mathcal{L} \). Let us define

\[
J^*(x, T|\lambda) := \sup_{\pi \in \mathcal{P}} \mathbb{E} \left[ \int_0^T p(s) \cdot dN^\pi(s) \right].
\]  

(4.2.5)

where the notation reflects the fact that the optimization problem is solved “conditioned” on knowing the demand function \( \lambda \) at time \( t = 0 \).

Clearly the value of the full information optimization problem (4.2.5) serves as an upper bound on the value of the original optimization problem described in (4.2.4). That is, for any demand function \( \lambda \in \mathcal{L} \) we have that \( J^\pi(x, T; \lambda) / J^*(x, T|\lambda) \leq 1 \) for all \( \pi \in \mathcal{P} \). This ratio measures the performance of any admissible policy on relative scale; generated revenues are expressed as a fraction of the optimal revenues in the full information setting. Our objective is to design policies that maximize this ratio uniformly over all demand functions in the class \( \mathcal{L} \); that is, choose \( \pi \in \mathcal{P} \) to maximize

\[
\inf_{\lambda \in \mathcal{L}} \frac{J^\pi(x, T; \lambda)}{J^*(x, T|\lambda)}.
\]  

(4.2.6)

The criterion in (4.2.6) can be viewed as the result of a two step procedure: first the decision maker selects a policy \( \pi \in \mathcal{P} \), and then “nature” picks the worst possible demand function \( \lambda \in \mathcal{L} \) for this particular policy. Measuring performance in this manner guarantees that “good” policies will perform well regardless of the
true underlying demand function. The fact that admissible policies can only learn
the true demand function by observing realized demand over time introduces an
obvious tension between exploration (estimation/demand learning) and exploita-
tion (optimization/pricing), and balancing these contradicting objectives is one of
the main issues that will be explored in what follows.

4.3 The Price Restricted Case

As alluded to earlier, the simplest instance of the blind network revenue man-
agement problem occurs when the set of feasible prices is discrete and finite, say,
\( P = \{p_1, \ldots, p_k, p_\infty\} \). This will be referred to as the price-restricted case in what
follows. In this setting, uncertainty is essentially limited to the value of the de-
mand function at the finite collection of prices in \( P \setminus \{p_\infty\} \). The main issue then
is how to estimate the mean demand rate at each price (exploration), and how
these prices will be used to extract maximal revenues (exploitation).

4.3.1 The proposed pricing policy

Our proposed blind pricing policy is based on a single tuning parameter \( \tau \in (0, T] \),
and is fleshed out in the pseudo-code given in the algorithm below; intuition is
discussed immediately thereafter.

Algorithm 5: \( \pi(\tau) \)

Step 1. Initialization:

Set the learning interval to be \([0, \tau]\) and put \( \Delta = \tau/k \).
Step 2. Learning/experimentation:

(a) While inventory is positive for all resources, price at $p_i$ from $t_{i-1} = (i - 1)\Delta$ to $t_i = i\Delta$, $i = 1, 2, ..., k$

If some resource is out of stock, apply $p_\infty$ up until time $T$ and STOP.

(b) Compute

\[ \hat{d}(p_i) = \frac{\text{total demand over } [t_{i-1}, t_i]}{\Delta}, \quad i = 1, ..., k. \]  

(4.3.1)

Step 3. Optimization/exploitation: Let $\hat{\hat{t}} = (\hat{t}_1, ..., \hat{t}_k)$ be the solution of the linear program

\[
\max \left\{ \sum_{i=1}^{k} p_i \cdot \hat{d}_i \hat{t}_i : \sum_{i=1}^{k} A\hat{d}_i \hat{t}_i \leq x, \sum_{i=1}^{k} \hat{t}_i \leq T - \tau, \hat{t}_i \geq 0, \; i = 1, ..., k \right\}. \tag{4.3.2}
\]

For each $i = 1, ..., k$, apply $p_i$ for $\hat{t}_i$ time units on $(\tau, T]$ until some resource is out of stock, then apply $p_\infty$ for the remaining time.

The intuition underlying the above construction is as follows. In Steps 1 and 2, the decision-maker estimates the demand at each of the $k$ feasible prices by testing the price on a period of time of length $\tau/k$. These estimates are then used in Step 3 to formulate an optimization problem, which is essentially an empirical version of the deterministic relaxation of the full information problem:

\[
\max \left\{ \sum_{i=1}^{k} p_i \cdot \lambda(p_i) t_i : \sum_{i=1}^{k} A\lambda(p_i) t_i \leq x, \sum_{i=1}^{k} t_i \leq T, \; t_i \geq 0, \; i = 1, ..., k \right\}. \tag{4.3.3}
\]
It is possible to show (see Gallego and van Ryzin (1997, 1994)) that the solution of this linear program leads to near-optimal performance in the full information stochastic dynamic optimization problem. The objective of Step 3 is to get “close” to said solution. As \( \tau \) increases, so do the quality of the estimates of the demand function values and in turn the quality of the approximation to the deterministic relaxation (and its solution). However, an increase in \( \tau \) also implies shorter exploitation time and higher potential revenue losses. This highlights the main trade-off faced in setting the value of \( \tau \).

### 4.3.2 Theoretical analysis

Exact analysis of the performance of the policy described in the previous section is quite difficult. We therefore introduce an asymptotic regime which facilitates an approximate analysis, and which has been used in several revenue management studies to date (see, e.g., Talluri and van Ryzin (2005, §3.6,5.3) and references therein). The regime is predicated on the number of initial resources and potential demand growing proportionally large. In particular, for any positive integer \( n \) the initial resource vector and the demand function are given by

\[
x_n = nx, \quad \lambda_n(\cdot) = n\lambda(\cdot).
\]

Here \( n \) which serves as a proxy for the market size determines both the order of magnitude of inventories and the rate of demand; when \( n \) is large this scaling characterizes a regime with a high volume of sales but maintains inventory constraints. The following notation will be useful: for real valued positive sequences \( \{a_n\} \) and
We write \( a_n = O(b_n) \) if \( a_n/b_n \) is bounded from above for large enough values of \( n \) (i.e., \( \limsup a_n/b_n < \infty \)). If \( a_n/b_n \) is also eventually bounded away from zero (i.e., \( \liminf a_n/b_n > 0 \)) then we write \( a_n \preceq b_n \).

We will denote by \( \mathcal{P}_n \) the set of admissible policies for a system with scale \( n \), and the expected revenues under a policy \( \pi_n \in \mathcal{P}_n \) will be denoted \( J_n^*(x,T;\lambda) \). With some abuse of notation we will occasionally use \( \pi \) to denote a sequence \( \{\pi_n : n = 1, 2, \ldots\} \) as well as any element of the sequence, omitting the subscript “\( n \)” to avoid clumping the notation. For each \( n = 1, 2, \ldots \), let \( J_n^*(x,T|\lambda) \) denote the optimal revenues that can be achieved in the full information case, i.e., when the demand function is known a priori in a system of scale \( n \). It follows from Section 2.2 that for all \( n = 1, 2, \ldots \), we have that \( J_n^*(x,T;\lambda) \leq J_n^*(x,T|\lambda) \). With this in mind, the following definition characterizes admissible policies that have “good” asymptotic properties.

**Definition 2 (Asymptotic optimality)** A sequence of admissible policies \( \{\pi_n\} \) is said to be asymptotically optimal if

\[
\inf_{\lambda \in \mathcal{L}} \frac{J_n^*(x,T;\lambda)}{J_n^*(x,T|\lambda)} \to 1 \quad \text{as} \quad n \to \infty.
\] (4.3.5)

Asymptotically optimal policies are those that achieve the full information upper bound on revenues as \( n \to \infty \), uniformly over the class of admissible demand functions. To that end, we have the following result.

**Theorem 4** For \( \pi_n \preceq n^{-1/3} \), the sequence of policies \( \{\pi(\tau_n)\} \) defined by Algorithm
5 is asymptotically optimal. In particular,

$$\sup_{\lambda \in \mathcal{L}} \left( 1 - \frac{J_n^*(x, T; \lambda)}{J_n(x, T|\lambda)} \right) = O\left( \frac{(\log n)^{1/2}}{n^{1/3}} \right) \quad \text{as } n \to \infty. \quad (4.3.6)$$

The right-hand-side above measures the revenue loss associated with using the proposed blind pricing policy, relative to the best achievable revenues under full information. As the theorem states, this gap in revenues shrinks to zero as the sales volume grows large. It is important to note that the rate at which this occurs does not depend on the number of products being sold.

### 4.3.3 A numerical illustration

We consider an example with two products and three resources. The first, second and third rows of the capacity consumption matrix $A$ are given by $(1,1)$, $(3,1)$ and $(0,5)$ respectively. This means that product 1 requires 1 unit of resource 1, 3 units of resource 2 and no units of resource 3, etc. We consider three different underlying demand models to test the efficacy of our proposed policy: a linear, an exponential and a logit model.

a) $\lambda(p_1, p_2) = (8 - 1.5p_1, 9 - 3p_2)'$,  
b) $\lambda(p_1, p_2) = (5\exp{-0.5p_1}, 9\exp{-p_2})'$,  
c) $\lambda(p_1, p_2) = 10(1 + \exp{-p_1} + \exp{-p_2})^{-1}(\exp{-p_1}, \exp{-p_2})'$.  

It is important to emphasize that our policies are constructed in a blind manner, without knowledge of the demand function. The set of feasible prices is $\{(1, 1.5), (1, 2), (2, 3), (4, 4), (4, 6.5)\}$. In Table 4.1, we illustrate the performance
of the policies defined by Algorithm 5 with $\tau_n = n^{-1/3}$. Note that in assessing the performance ratio $J_n^/ / J_n^*$, we use the upper bound provided by the deterministic relaxation in place of $J_n^*$ (see Gallego and van Ryzin (1997)) and hence the actual ratio $J_n^/ / J_n^*$ is at least as high as that reported in the table. The results are based on running $10^3$ independent simulation replications from which the performance indicators were derived by averaging. The standard error was below 0.1% in all cases.

<table>
<thead>
<tr>
<th>Market &quot;size&quot;</th>
<th>$n = 10^2$</th>
<th>$n = 10^3$</th>
<th>$n = 10^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuning parameters</td>
<td>$\tau = 0.22$</td>
<td>$\tau = 0.10$</td>
<td>$\tau = 0.05$</td>
</tr>
<tr>
<td>$J_n^/ / J_n^*$</td>
<td>$J_n^/ / J_n^*$</td>
<td>$J_n^/ / J_n^*$</td>
<td></td>
</tr>
<tr>
<td>initial inventory</td>
<td>Linear</td>
<td>.65</td>
<td>.86</td>
</tr>
<tr>
<td>$x_n = (3, 5, 7) \times n$</td>
<td>Exponential</td>
<td>.75</td>
<td>.84</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Logit</td>
<td>.78</td>
<td>.87</td>
</tr>
<tr>
<td>$x_n = (15, 12, 30) \times n$</td>
<td>Linear</td>
<td>.76</td>
<td>.83</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Exponential</td>
<td>.87</td>
<td>.94</td>
</tr>
<tr>
<td>$x_n = (15, 12, 30) \times n$</td>
<td>Logit</td>
<td>.88</td>
<td>.94</td>
</tr>
</tbody>
</table>

Table 4.1: **Price restricted case.** Results in the table give a lower bound on the performance ratio of the policy $\pi(\tau)$ to the optimal performance in the full information case. Here $\tau = \text{fraction of time allocated to learning}$.

The results in Table 4.1 are consistent with the asymptotic optimality statement of Theorem 4. The proposed policy generates at least 83% of the full information benchmark for $n = 10^3$, and this performance is achieved by allocating only 10% of the selling horizon to the learning phase. In addition, the performance is seen to be comparable for the various demand models tested, exhibiting the robustness asserted in Theorem 4.1.
4.4 The General Case

In the price restricted case studied in the previous section, demand model uncertainty manifested itself in only a finite number of relevant values. When firms are free to select their prices beyond a finite collection, the feasible price set \(D_p\) becomes uncountable and model uncertainty pertains to a surface \(\lambda(\cdot)\) in a \(d\)-dimensional space. We henceforth assume that \(D_p\) is a compact convex set, and one of the key questions in the design of blind pricing policies is now concerned with the selection of "good" test prices within the feasible set.

4.4.1 The blind pricing policy

Before describing the policy we need to define a price grid. Let \(B_p := \prod_{i=1}^{d}[p^i, \bar{p}^i]\) denote the minimum volume hyper-rectangle in \(\mathbb{R}^d_+\), such that \(B_p \supset D_p\). Given a positive integer \(\kappa\), one can divide each interval \([p^i, \bar{p}^i], i = 1, \ldots, d\) into \([\kappa^{1/d}]\) intervals of equal length. Define the resulting grid of points in \(\mathbb{R}^d_+\) as \(B^\kappa_p\). Let \(e = (1, \ldots, 1) \in \mathbb{R}^d\). The following algorithm provides pseudo-code that defines a class of admissible learning and pricing policies that are parametrized by a triplet of tuning parameters \((\tau, \kappa, \delta)\): \(\tau \in (0, T]\) represents the length of an initial interval dedicated to learning; \(\kappa\) is a positive integer that defines the number of prices to "test" during the learning phase; and \(\delta > 0\) is a "fudge factor" that allows for some slack in the capacity constraint (4.2.2).

Algorithm 6: \(\pi(\tau, \kappa, \delta)\)
Step 1. Initialization:

(a) Set the learning interval to be $[0, \tau]$, and set $\kappa$ to be the number of prices to experiment with. Put $\Delta = \tau / \kappa$.

(b) Define $P^K = \{p_1, \ldots, p_{\kappa}\}$ to be the prices to experiment with over $[0, \tau]$, where $P^K \supseteq B^\kappa \cap D_p$.

Step 2. Learning/experimentation:

(a) On the interval $[0, \tau]$ apply $p_i$ from $t_{i-1} = (i-1)\Delta$ to $t_i = i\Delta$, $i = 1, 2, \ldots, \kappa$ as long as inventory is positive for all resources. If some resource is out of stock, apply $p_\infty$ up until time $T$ and STOP.

(b) Compute

$$\hat{d}(p_i) = \frac{\text{total demand over } [t_{i-1}, t_i]}{\Delta}, \quad i = 1, \ldots, \kappa. \quad (4.4.1)$$

Step 3. Optimization:

For $i = 1, \ldots, \kappa$,

If $A\hat{d}(p_i)T \leq x + \delta e$, then [check if price is feasible]

$$\hat{r}(p_i) = p_i \cdot \hat{d}(p_i) \quad \text{[compute empirical revenue rate]}$$

else $\hat{r}(p_i) = 0$.

End If

End For
Set \( \hat{p} = \arg\max \{ \hat{r}(p) : p \in P^K \} \).  

[empirically optimal price]

**Step 4. Pricing:**

On the interval \((r, T]\) apply \( \hat{p} \) until some resource is out of stock, then apply \( p_\infty \) for the remaining time.

In Step 3, \( A \) denotes the capacity-consumption matrix defined in the capacity constraint (4.2.2) of the original dynamic optimization problem (Section 2.2). Regarding Step 4, it is clear that any practical implementation of the policy would not “shut off” *all* the demand once a single resource becomes unavailable, but would rather do so only for those products that use the unavailable resource. The result we present in Theorem 5 is valid for policies that improve upon the above by refining Step 4 through partial and/or gradual demand “shut off.”

**Intuition.** Step 1 sets the first two tuning parameters: \( r \) determines the length of interval used for learning the demand function; and \( K \) sets the number of prices that are experimented with on that interval. In Step 2, prices in the discret set \( P^K \) are used to obtain an empirical approximation of the demand function;

To understand the logic underlying Step 3, imagine that the demand function \( \lambda(\cdot) \) is revealed at the start of the selling season, and demand is deterministic rather than governed by a Poisson process. The revenue maximization problem
would then be given by the following deterministic dynamic optimization problem

$$\max \left\{ \int_0^T r(\lambda(p(s)))ds : \int_0^T A\lambda(p(s))ds \leq x, \ p(s) \in D_p \text{ for all } s \in [0,T] \right\},$$

(4.4.2)

where $r(\cdot)$ is the revenue rate. Gallego and van Ryzin (1997) show that the solution to (4.4.2) is constant over time, and establish that this fixed price yields close to optimal performance in the original stochastic problem. Step 3 of the algorithm uses observed demand to form an estimate of the revenue function, and then proceeds to solve a suitable empirical version of the deterministic problem (4.4.2).

The optimal solution for this problem $\hat{p}$ is then used for the remainder of the time horizon $(\tau, T]$. The choice of the tuning parameter $\delta$ allows some modest violation of the capacity constraints: the logic here is that the estimates of the demand rate are "noisy," and the $\delta$-slack avoids restricting too drastically the search for the empirical optimal price $\hat{p}$.

**Balancing the exploration-exploitation trade-off.** The choice of the key tuning parameters, the learning horizon $\tau$ and number of "test prices" $\kappa$ is meant to balance several contradicting objectives. Specifically, increasing $\tau$ results in a longer time horizon over which the demand function is estimated, however by doing so there is also a potential loss in revenues that stems from spending "too much time" on learning and exploration. For every fixed choice of $\tau$, there is an inherent tradeoff between the number of prices to experiment with, $\kappa$, and the accuracy of estimating the demand function on this price grid which is dictated by the length $\Delta = \tau/\kappa$. In particular, using more prices translates into a better coverage of the
domain of the demand surface, but it also implies that the estimates are more “noisy” since each price is used for a shorter interval. The next section explains how to balance these error sources.

### 4.4.2 Performance analysis

In addition to the basic assumptions outlined in Section 2.2, we impose the following regularity conditions, which are quite standard in the revenue management literature (cf. Talluri and van Ryzin (2005)).

**Assumption 8** Every demand function \( \lambda(\cdot) \in \mathcal{L} \) has an inverse, denoted \( \gamma(\cdot) \), the set \( \mathcal{D}_\lambda := \{ l : l = \lambda(p), p \in D_p \} \) is convex and the revenue function \( r(\lambda) := \lambda \cdot \gamma(\lambda) \) is concave. In addition, \( \lambda(\cdot) \) is Lipschitz continuous, i.e., \( \| \lambda(p) - \lambda(p') \| \leq K \| p - p' \| \) for all \( p, p' \in D_p \), where \( K \) is a given positive constant.

In the context of the high sales volume asymptotic regime given in (4.3.4), we provide below a result that characterizes the performance of blind policies defined by Algorithm 6.

**Theorem 5** Let Assumption 8 hold, and set

\[
\tau_n \asymp n^{-1/(d+3)}, \quad \kappa_n \asymp n^{d/(d+3)}, \quad \delta_n = C n (\log n)^{1/2} n^{1/(d+3)},
\]

with \( C > 0 \) sufficiently large. Then the sequence of policies \( \{ \pi_n \} \) defined by Algorithm 6 is asymptotically optimal. In particular,

\[
\sup_{\lambda \in \mathcal{L}} \left( 1 - \frac{J_n^x(x, T; \lambda)}{J_n^x(x, T|\lambda)} \right) = O\left( \frac{(\log n)^{1/2}}{n^{1/(d+3)}} \right) \quad \text{as } n \to \infty,
\]

where \( d \) denotes the number of products.
Remarks.

1. As in Theorem 4, the first part of the theorem states that the value of full information diminishes for large $n$. Note that the length of the exploration phase $\tau_n$ shrinks as $n$ gets large, and in that sense learning occurs on a shorter time scale than the sales horizon $[0, T]$. In particular, we can rewrite (4.4.4) informally (ignoring logarithmic terms) as $J_n^x(x, T; \lambda)/J_n^*(x, T|\lambda) \approx 1 - C\tau_n/T$ and hence the "price" (revenue loss) of not knowing the demand function a priori is proportional to the relative size of the learning horizon.

2. It is important to note that, in contrast with the price-restricted case and the rate of revenue loss provided in (4.3.6), in this setting this rate degrades with the number of products $(d)$; an obvious manifestation of the curse of dimensionality. We return to this point in Section 4.5, where a method is proposed to diminish this effect.

Proof sketch. As alluded to in the discussion following Algorithm 5, there are three sources of error that impact the revenue loss relative to the maximal full information revenue benchmark, as captured by the ratio $J_n^x/J_n^*$. The first error source can be interpreted as an "exploration bias" that is due to experimenting with various prices in the absence of any information on the demand model. This will result in potential revenue losses of order $\tau_n$. The second error source is deterministic and stems from restricting to only a finite number of prices the search for the optimal solution of (4.4.2); the maximal loss related to this error is of order $1/\kappa_n^{1/d}$ which is the granularity of the price grid. The last source of error is
stochastic, arising from the fact that only noisy observations of the demand function are available. Since each price is held fixed for \( \Delta_n = \tau_n / \kappa_n \) units of time, this introduces an error of order \((n\tau_n / \kappa_n)^{-1/2}\); this observation is less transparent and is rigorously detailed in the proof using uniform probability bounds for deviations of random variables from their expectation. The overall revenue loss is dictated by the sum of the three sources detailed above, namely

\[
1 - J_n^* / J_n^* \approx \mathcal{O}\left(\tau_n + \frac{1}{\kappa_n^{1/2}} + \frac{\kappa_n^{1/2}}{(n\tau_n)^{1/2}}\right).
\] (4.4.5)

This last expression captures mathematically the tension that must be resolved in choosing the tuning parameters associated with Algorithm 6. Roughly speaking, shortening \( \tau_n \) decreases the exploration bias, but increases the stochastic error since there is more "noise" at each tested price. Similarly, increasing the number of test prices \( \kappa_n \) shrinks the deterministic error, but increases the stochastic error since more prices need to be tested in \( \tau_n \) time units. Balancing the three error terms in (4.4.5) yields the choice of tuning parameters reported in the theorem and gives rise to the revenue loss rate in (4.4.4).

### 4.4.3 A simple state-dependent refinement

The learning phase in the policy described by Algorithm 6 results in an estimate of the demand function at the price vectors that are tested over the interval \([0, \tau)\). These estimates are subsequently used to solve an empirical version of the full information deterministic relaxation problem, which results in a single price which is then used for the rest of the selling season. This strategy does not make further
use of the estimates of the demand function after time $t = \tau$. A simple way to refine this approach would be to re-solve the aforementioned optimization problem at additional points in time downstream of $\tau$. For example, consider a policy $\pi^*_2(\tau, \kappa, \delta, T_r)$ that re-solves at time $T_r$. It proceeds as in Algorithm 6 except that Step 4 is replaced by:

---

**Step 4$^{(r)}$. Pricing:**

On the interval $(\tau, T_r]$ apply $\hat{p}$. If some resource is out of stock, apply $p_{\infty}$ up until time $T$ and STOP. Otherwise, let $I_r$ be the inventory at time $T_r$ and re-solve:

For $i = 1, \ldots, \kappa$,

If $A\tilde{d}(p_i)(T - T_r) \leq I_r + \delta e$, then

$$\hat{r}^{(2)}(p_i) = p_i \cdot \tilde{d}(p_i)$$

else $\hat{r}^{(2)}(p_i) = 0$.

End If

End For

Set $\hat{p}^{(2)} = \text{arg max}\{\hat{r}^{(2)}(p) : p \in P^\kappa\}$.

On the interval $(T_r, T]$ apply $\hat{p}^{(2)}$ until some resource is out of stock, then apply $p_{\infty}$.
Intuition. While such re-solving strategies are not guaranteed to yield benefits (see, e.g., Cooper (2002) in the context of a capacity allocation problem), the main idea here is to allow for some adaptation of the price to a given sample path of demand. As our discussion following Theorem 5 indicates, the average performance of the policy described in Algorithm 6 is essentially dictated by a law of large numbers, hence introducing re-solving points is expected to “hedge” against deviations from the average case behavior, and lead to potential improvements in performance. We illustrate this in the next section.

4.4.4 Illustrative numerical examples

Note that, as in the price restricted case, \( J_n^*(x,T|\lambda) \) is not readily computable in most cases. However, an upper bound is easy to obtain through the value of the deterministic optimization problem given in (4.4.2). This upper bound is fairly tight for moderate sized problems (see Gallego and van Ryzin (1997)), and hence one can compute a “good” lower bound on the ratio \( J_n^*(x,T;\lambda)/J_n^*(x,T|\lambda) \) based on this deterministic relaxation. The results depicted in Table 4.2 were obtained by running \( 10^3 \) independent simulation replications from which the performance indicators were derived by averaging. The standard error was below 0.5% in all cases.

The capacity consumption matrix \( A \) and true demand functions are the ones defined in Section 4.3.3. The set of feasible prices is taken to be \( \mathcal{D}_p = [0.5, 5] \times \).
[0.5, 5] and $T = 1$. In Table 4.2, we give performance results for the policy $\pi$ defined by Algorithm 6, and the re-solving policy $\pi^r$ (with tuning parameters given in (4.4.3), with $C = 2$ and $T_r = 1/2$ for each $n$).

<table>
<thead>
<tr>
<th>Market &quot;size&quot;</th>
<th>$n = 10^2$</th>
<th>$n = 10^3$</th>
<th>$n = 10^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tune parameters</td>
<td>$\kappa = 9$, $\tau = 0.29$</td>
<td>$\kappa = 16$, $\tau = 0.18$</td>
<td>$\kappa = 36$, $\tau = 0.12$</td>
</tr>
<tr>
<td>Policy</td>
<td>$\pi$</td>
<td>$\pi^r$</td>
<td>$\pi$</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Linear</td>
<td>.57</td>
<td>.60</td>
</tr>
<tr>
<td>$x = (3, 5, 7)$</td>
<td>Exp</td>
<td>.73</td>
<td>.75</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Logit</td>
<td>.68</td>
<td>.67</td>
</tr>
<tr>
<td>$x = (15, 12, 30)$</td>
<td>Linear</td>
<td>.77</td>
<td>.78</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Exp</td>
<td>.69</td>
<td>.68</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Logit</td>
<td>.68</td>
<td>.68</td>
</tr>
</tbody>
</table>

Table 4.2: Results in the table give a lower bound on the ratio of the revenues extracted by the policies $\pi(\tau, \kappa, \delta)$ (Algorithm 6) and $\pi^r(\tau, \kappa, \delta, T_r)$ (Algorithm 6 with Step 4($r$)), to the optimal revenues in the full information case. Here $\kappa =$ number of prices tested by the policy, $\tau =$ fraction of time allocated to learning, and $T_r = 1/2$ is the time of re-solving and price adjustments in $\pi^r$.

We observe that with inventory levels of the order of a few thousands, the expected revenues under the proposed policy $\pi_2$ exceed 65% of the optimal expected revenues in the full information case (where the demand function is known a priori). The policy utilizes approximately 18% of the time horizon $T$ to learn the demand function and experiments with 16 prices. Inspecting the results, we observe that the ratio $J_n(\pi)/J_n^*$ approaches 1 as the market size increases, as predicted by the asymptotic optimality result in Theroem 5. We also observe that the performance of the re-solving policy $\pi^r_2$ is roughly on par with that of the original policy $\pi_2$ and does not yield significant improvements. This suggests that dynamic price adjustments following the learning phase have little impact on performance and that the latter is primarily driven by the uncertainty associated with the demand.
function. Comparing the results above with those in Table 4.1 that summarize the price restricted case, it is evident that superior performance is observed when the set of feasible prices is finite; this is in line with intuition and the performance guarantees given in Theorems 4 and 5, respectively. In particular, the performance in the price-restricted case does not degrade with the dimension of the set of products.

4.5 The Curse of Dimensionality

The policies we have outlined in Section 4.4 suffer from the curse of dimensionality, which leads to a degradation in the performance guarantee given in Theorem 5 (as the number of products $d$ increases). The culprit here is the necessity to experiment with sufficiently many price combinations to “cover” the domain of the unknown demand function. To mitigate this problem, and improve the result in (4.4.4), one would need to exploit additional structural information on the unknown demand function. We first explore how one can exploit further smoothness of the demand function; such information can be used to reconstruct the entire demand surface through techniques from nonparametric statistics and to design more “clever” pricing policies. Then we comment briefly on the case where the demand function is separable (a common assumption in the literature).

4.5.1 Exploiting smoothness of the demand function

Rather than restricting smoothness to the function, as in Assumption 8, we now essentially require that the Lipschitz condition given there hold for the first $s - 1$
derivatives. Thus, our class of demand functions is assumed to be $s$-times differentiable with uniformly bounded derivatives. Note that almost all demand functions commonly used in the literature fall into this category. We state this more formally in the following assumption.

**Assumption 9** For some constant $L > 0$ and positive integer $s$, the demand function $\lambda$ is $s$ times differentiable, and for all $i = 1, \ldots, d$

$$\left| \frac{\partial^{a_1, \ldots, a_d} \lambda_i(p)}{\partial p_1^{a_1} \cdots \partial p_d^{a_d}} \right| \leq L$$  (4.5.1)

for all $p \in D_p$ and nonnegative integers $a_1, \ldots, a_d$ such that $a_1 + \ldots + a_d = s$.

Here $d$ is the dimension of the set of products, and $s$ is mnemonic for smoothness of the demand function. The idea now is the following: given a price grid, e.g., as detailed prior to the statement of Algorithm 6, and the observed demand at each of those price vectors $y = (\hat{d}(p_1), \ldots, \hat{d}(p_\kappa))$, reconstruct an approximation to the entire demand surface $\hat{\lambda}(p; y)$ over the price domain, $p \in D_p$. To achieve this goal, we will focus here on a nonparametric method based on local polynomials, which roughly works as follows: for a given price point $p \in D_p$, consider a neighborhood of that point $B_p$ which is a hypercube with edge length $h^1$; fit a polynomial of degree $s - 1$ to that neighborhood using observed demand and approximate the value of the function $\lambda(\cdot)$ at $p$ by that of the polynomial at the same point.

More specifically, let us focus initially on the first component of the demand function $\lambda^1(\cdot)$ and detail the development of the approximation. For $i = 1, \ldots, \kappa$

\[ B_p = \prod_{i=1}^d B_i \text{ where } B_i = \begin{cases} [p_i, p_i + h] & \text{if } p_i \leq p_i^* + h/2, \\ [p_i^* - h, p_i^*] & \text{if } p_i \geq p_i^* - h/2 \text{ and } B_i = [p_i^* - h/2, p_i^* + h/2] & \text{otherwise} \end{cases} \]
and \( j = 1, \ldots, d \), \( y^j_i \) denotes the number of requests for product \( j \) when pricing at \( p^i \) in the learning phase, \( y^j \) denotes the row vector \( (y^j_1, \ldots, y^j_N) \), and \( y_i \) denotes the column vector \( (y^1_i, \ldots, y^d_i)^T \). Select the parameter \( h > 0 \) such that \( h\kappa^{1/d} \geq s + 1 \).

For every \( p \in \mathcal{D}_p \), we define a window \( B_p = \prod_{i=1}^d B^i \), where \( B^i = [p^i, p^i + h] \) if \( p^i \leq p^i + h/2 \), \( B^i = [p^i - h, 1] \) if \( p^i \geq p^i - h/2 \) and \( B^i = [p^i - h/2, p^i + h/2] \) otherwise.

The local polynomial approximation to the function \( \lambda^1(\cdot) \) will be a weighted sum of the observations \( y^1_i \). We construct a set of weights as follows. Let \( \beta^1, \ldots, \beta^N \) be a basis in the space of polynomials of degree \( s - 1 \) of \( d \) variables. Fix a price \( p \in \mathcal{D}_p \) and denote by \( G = B_p \cap P^\kappa \) the set of price points in the grid that also lie in the window \( B_p \). Let \( \beta^i_G \) denote the column vector whose \( j^{th} \) component is given by the value of \( \beta^i \) at the \( j^{th} \) point in \( G \) amd let \( M = [\beta^1_G, \ldots, \beta^N_G] \). Given that \( h\kappa^{1/d} \geq s + 1 \), it is possible to show that \( M \) has full rank (see Nemirovski (2000)). We define a vector of weights \( \omega^B(p) \) as follows

\[
\omega^B(p) = M(M^TM)^{-1}V(p),
\]

where \( V(p) = (\beta^1(p), \ldots, \beta^N(p))^T \). Given the weights, the approximation takes the following form

\[
\hat{\lambda}^1(p; y^1) = \sum_{i:p_i \in B_p} \omega^B_i(p) y^1_i,
\]

A similar approach conducted for every component of the demand function yields the approximation

\[
\hat{\lambda}(p; y) = \sum_{i:p_i \in B_p} \omega^B_i(p) y_i.
\]
The proof of Theorem 6 in Appendix C.1 contains further details on the approximation and some of its properties. Nemirovski (2000) is a recent reference on such approximations.

To describe a policy that uses this nonparametric regression methodology, consider now replacing Step 3 of Algorithm 6 by the following Step 3':

---

Step 3'. Optimization:

a) Let $y_i = d(p_i)$, $i = 1, ..., \kappa$

b) Let $\hat{\lambda}(\cdot, y)$ be an approximation to $\lambda(\cdot)$ based on local polynomials of order $s - 1$

c) Set $\hat{\rho} = \arg \max_{p \in D_{r}} \{p \cdot \hat{\lambda}(p; y) : A\hat{\lambda}(p; y) \leq x + \delta e\}$

---

Let Algorithm 6' denote Algorithm 6 where Step 3' replaces Step 3. The policy described by Algorithm 6' takes as input four tuning parameters $(\tau, \kappa, \delta, h)$, where $h$ is a smoothing parameter associated with the local polynomial regression to estimate and reconstruct the demand function; See footnote 1 and further details in the proof of Theorem 6. In the context of the asymptotic regime given in (4.3.4), the performance of policies defined by means of Algorithm 6' is given in the following result.

**Theorem 6** Let Assumptions 8 and 9 hold. Let $\pi$ denote the policies defined by Algorithm 6' where $\tau_n \asymp n^{-1/(3+d/s)}$, $\kappa_n \asymp \lceil n^{d/(3s+d)} \rceil$, $h_n = (s + 1)^{-1} \kappa_n^{-1/d}$, and
\( \delta_n \preceq C(\log n)^{1/2} n^{-1/(3+d/s)} \) with \( C > 0 \) sufficiently large, then \( \{\pi_n\} \) is asymptotically optimal and

\[
\sup_{\lambda \in \mathcal{L}} \left( 1 - \frac{J_n(x, T; \lambda)}{J^*(x, T|\lambda)} \right) = O\left( \frac{(\log n)^{1/2}}{n^{1/(3+d/s)}} \right). \tag{4.5.5}
\]

**Remark 1 (the curse of dimensionality).** While the revenue loss relative to the full information optimal revenues given in (4.5.5) degrades as the number of products, \( d \), increases, it is now evident that the smoother the demand function, the lesser are the curse of dimensionality effects. In particular the rate of convergence for the policy given by Algorithm 6' is \( n^{-1/(3+d/s)} \) compared to \( n^{-1/(3+d)} \) for the original policy described by Algorithm 6. Note that if the demand function is “very smooth” (roughly speaking infinitely continuously differentiable), then \( J_n^n/J_n^* \approx 1 - C/n^{1/3} \), up to logarithmic terms. That is, revenue losses resulting from not knowing the demand function are *dimension-independent*, approaching the performance of the price restricted case given in Theorem 4. Given that most commonly used demand functions are very smooth, the potential of our proposed approach is evident.

**Remark 2 (implications for price testing).** Theorem 6 implies that the number of prices to be tested in the learning phase is reduced compared to Theorem 6. This is an important implication from a practical perspective: Algorithm 6' exploits smoothness of the demand function to extract more information (per tested price). The theoretical basis for this can be found in Theorems 5 and 6 if one focuses on the magnitude of the number of price tests: \( \kappa_n = n^{d/(d+3)} \) in the former;
and \( n^{d/(d+3s)} \) in the latter. The intuition here is that a smoother function can be “pinned down” using fewer points.

**Intuition.** The main intuition underlying the result in Theorem 6 is as follows: as the smoothness of the underlying demand function increases, the variation of this function between any two points becomes more and more restricted. Exploiting this yields an improvement in the approximation of the demand function. In particular, one can show that with tuning parameters \( \kappa_n \) and \( h_n \) chosen such that \( \kappa_n \asymp \left[n^{d/(3s+d)}\right] \) and \( h_n = (s+1)^{-1} \kappa_n^{-1/d} \), one has

\[
\sup_{\lambda \in \mathcal{L}} \mathbb{E} \| \hat{\lambda}_n(p;y) - \lambda(p) \|_\infty \approx (n\tau_n)^{-\frac{s}{2s+d}}. \tag{4.5.6}
\]

Let us revisit the proof sketch of Theorem 5 (outlined following the statement of that result). Three main error sources were highlighted there: an exploration bias; a deterministic error; and a stochastic error. In the current context, the key observation is that the magnitude of the deterministic and stochastic errors can be reduced by exploiting smoothness. In particular, we now have for all \( \lambda \in \mathcal{L} \)

\[
1 - \frac{J_n^\pi(x,T;\lambda)}{J_n^*(x,T;\lambda)} \approx \tau_n + (n\tau_n)^{-\frac{s}{2s+d}}, \tag{4.5.7}
\]

where \( s \) is the smoothness index and \( d \) is the dimensionality. Balancing the error sources, one gets that the optimal choice of the learning horizon is \( \tau_n \approx n^{-s/(3s+d)} \).

Now, one has that the fraction of the optimal full information revenue extracted by \( \pi \) is of order \( J_n^\pi / J_n^* \approx 1 - \tau_n \). The rate \( \tau_n \) degrades gracefully with the dimension \( d \), due to the smoothness of the demand function which is exploited by the policy.
A numerical example: performance analysis. To illustrate the performance of Algorithm 6' defined above, we consider the same setting as in Section 4.4.4. Let \( \pi \) denote the policy given by Algorithm 6. Let \( \pi' \) denote the policy that follows Algorithm 6' and uses local polynomials of degree 1 to approximate the demand function. In Table 4.3, we provide performance results when both policies use the same tuning parameters. This comparison highlights the value of "reconstructing" the demand function (as in Step 3' of Algorithm 6') rather than restricting the search to prices that were tested in the learning phase (as in Step 3 of Algorithm 6).

<table>
<thead>
<tr>
<th>Market &quot;size&quot;</th>
<th>( n = 10^3 )</th>
<th>( n = 10^4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuning parameters</td>
<td>( \kappa = 16, \tau = 0.18 )</td>
<td>( \kappa = 36, \tau = 0.12 )</td>
</tr>
<tr>
<td>Policy</td>
<td>( \pi )</td>
<td>( \pi' )</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Linear</td>
<td>.65</td>
</tr>
<tr>
<td>( x_n = (3, 5, 7) \times n )</td>
<td>Exponential</td>
<td>.77</td>
</tr>
<tr>
<td></td>
<td>Logit</td>
<td>.81</td>
</tr>
<tr>
<td>initial inventory</td>
<td>Linear</td>
<td>.85</td>
</tr>
<tr>
<td>( x_n = (15, 12, 30) \times n )</td>
<td>Exponential</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>Logit</td>
<td>.83</td>
</tr>
</tbody>
</table>

Table 4.3: Exploiting smoothness. Results in the table give a lower bound on the ratio of the revenues extracted by the policy \( \pi \) (Algorithm 6) and \( \pi' \) (Algorithm 6'), relative to the optimal revenues in the full information case. Here \( \kappa = \) number of prices tested by the policy, and \( \tau = \) fraction of time allocated to learning.

We observe a general improvement in performance when using policy \( \pi' \) compared to what is achieved by policy \( \pi \). This improvement can be significant, at times exceeding 10% of the full information optimal revenues. In particular, for the examples considered, the performance of policy \( \pi' \) always exceeds 76% of the
full information optimal revenues for market sizes of the order of $10^3$. The improvements are more marginal for cases where initial inventories are large (the case $x = (15, 12, 30)$); where $\pi$ already achieves a good performance. These results illustrate that it is possible to take advantage of the smoothness of the underlying unknown demand function to reconstruct an estimate of the entire demand function, and hence improve upon the revenue optimization. Note that the performance of policies $\pi$ and $\pi'$ was illustrated using identical tuning parameters, and in particular the same number of prices were tested. Thus, one may interpret the numerical results as follows: for any given performance level, fewer prices need to be tested when using $\pi'$ as opposed to $\pi$.

**A numerical example: reconstructing the demand function.** In order to illustrate how local polynomials are used to reconstruct the demand function and in turn the revenue function, we depict in Figure 4.1(a) the revenue function derived from a logit model $(\lambda(p_1, p_2) = 10(1+\exp{-p_1}+\exp{-p_2})^{-1}(\exp{-p_1}, \exp{-p_2}))$ and in Figure 4.1(b) the approximation obtained by using local polynomials of degree 1 that is used by the policy $\pi'$ with market size of $10^4$. Note that this approximation is based on a single realization of the learning phase. In particular, 36 prices are tested in the domain $\mathcal{D}_p = [0.5, 5] \times [0.5, 5]$, and the resulting demand observations are used to reconstruct the demand function.

If one focuses on the iso-revenue contours, one observes that the general shape of the revenue function is recovered reasonable well and in particular, the location of the maximizer of the revenue function is well approximated. This is one of
Figure 4.1: **Reconstructing the revenue surface.** (a) revenue function derived from a logit demand model; (b) approximation to the revenue function using local polynomials of degree 1. Iso-revenue contours are indicated on the \((p_1, p_2)\) plane. The construction is obtained by testing 36 prices.

the reasons why the policy \(\pi'\) works so well. The main takeaway here is that a relatively small number of prices allows for a reasonable accurate reconstruction of the demand function on the entire price domain, and this translates to the performance improvement reported in Table 4.3.

### 4.5.2 Exploiting separability of the demand function

If it is known a priori that the demand function is *separable*, meaning that the demand for the \(i^{th}\) product, \(\lambda^i(p)\), is only a function of the price of the \(i^{th}\) product
then the products are only linked through the capacity constraints in the revenue optimization problem. In such cases, one can refine the learning stage of Algorithm 6, by testing every product price once (since the demand is separable), as opposed to having to test all price combinations in the price grid. From this, one can construct estimates of the demand function at all price vectors in the price grid. By limiting the number of price tests as above, it is straightforward to show that the relative revenue loss would be

$$1 - \frac{J_n^*(x, T; \lambda)}{J_n^*(x, T; \lambda)} \approx n^{-1/4}. \quad (4.5.8)$$

In particular, performance no longer depends on the number of products that are being sold and hence is dimension-independent. The proof of this result is omitted as it follows exactly along the same lines as the proof of Theorem 5. (Of course, one can combine this observation with the method outlined earlier to exploit smoothness and further improve on (4.5.8)).

4.6 Concluding Remarks

Incorporating additional prior information. The policies proposed in this paper make no significant a priori assumptions on the unknown demand function or its domain (feasible price set). In particular, all prices within the domain $\mathcal{D}_p$ are effectively assumed to be equally likely candidates for generating the highest revenues in the deterministic relaxation. If certain price combinations are considered to be more likely to generate higher profit versus others, one can encode such beliefs in the form of a prior distribution with support contained in $\mathcal{D}_p$. The
structure of this prior would then lead to the use of non-uniform price grids in the proposed algorithms: regions in $D_p$ that have higher prior probability mass will be "quantized" more finely as opposed to other areas which have lower probability.

**On the Poisson process assumption.** We have made the assumption that requests for products arrive according to a Poisson process whose rate is given by the underlying demand function (evaluated at a given price). This assumption is made primarily for concreteness and in order to keep technical details to a bare minimum. In essence, the notion of asymptotic optimality we advocate in this paper only relies on a rough estimate of the rate of convergence in the strong law of large numbers. Thus, the results given in Sections 4.3-4.5 can be derived under far more general assumptions on the underlying point process that governs demand.

**Extensions.** Our approach hinges on the fact that the revenue management problem being discussed can be "well approximated" by an appropriate deterministic relaxation which admits a simple solution. This is encoded in Step 3 of the policies described in Sections 4.4 and 4.5 of this paper. Roughly speaking, this ensures that a static fixed price nearly maximizes revenues in the full information case (cf. Gallego and van Ryzin (1994, 1997)). Problems that admit such structure appear in various other contexts, see, e.g., Paschalidis and Tsitsiklis (2000) and Maglaras and Zeevi (2005)), and the techniques developed in this paper may prove
useful in those problems as well.

**Adaptive algorithms.** Even though the performance of the proposed policies was shown to be near-optimal, the decision maker will not necessarily wish to fully separate the learning phase from the pricing phase in the manner prescribed by Algorithms 5, 6 and 6'. In particular, it is possible to make the learning phase *adaptive* so that only relevant regions of the feasible price set are explored. That is, the estimation and optimization stages might be pursued simultaneously rather than sequentially, to perform a better localized search for the near optimal fixed price.
Bibliography


Carvalho, A. X. and Puterman, M. L. (2005), ‘Dynamic pricing and learning over short time horizons’, *working paper, University of British Columbia*.


Appendix A

Proofs for Chapter 2

A.1 Proof of the Main Result

Proof of Theorem 2. The proof is organized in three main steps. In a first step, we decompose $\Delta_n$ in two terms, whose behavior under $H_0$ and $H_1$ are subsequently analyzed separately in Lemmas 1 and 3.

We decompose $\Delta_n$ as follows

$$\Delta_n = A_n + B_n,$$

where

$$A_n = \hat{\pi}_n(\hat{x}_n) - \hat{\pi}_n(x^*)$$

$$B_n = \hat{\pi}_n(x^*) - \hat{\pi}_n(x^*(\hat{\theta}))$$

Next, we analyze each term $A_n$ and $B_n$ separately under the null and the alternative hypotheses.

Lemma 1 Suppose that $nh_n^6 \to \infty$ and $nh_n^7 \to 0$, then

$$nh_n^3 A_n \Rightarrow \gamma \chi^2,$$
where γ was defined in (2.5.4) and \( \chi^2 \) is a Chi-squared random variable with one degree of freedom.

**Proof of Lemma 1.** Noting that \( \hat{\pi}(\cdot) \) is differentiable, a Taylor expansion gives

that that for some \( x_{1,n} \in [\min\{\hat{x}_n, x^*\}, \max\{\hat{x}_n, x^*\}] \)

\[
A_n = -\hat{\pi}'(\hat{x}_n)(x^* - \hat{x}_n) - \frac{1}{2} \hat{\pi}''(x_{1,n})(x^* - \hat{x}_n)^2
\]

\[
= -\frac{1}{2} \hat{\pi}''(x_{1,n})(x^* - \hat{x}_n)^2 \quad \text{(A.1-1)}
\]

Now, by Ziegler (2002, Theorem 3.1), we have under the assumption that \( nh_n^6 \rightarrow +\infty \) and that \( nh_n^7 \rightarrow 0 \)

\[
\sqrt{nh_n^3(x^* - \hat{x}_n)} \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2(x^*)}{(\pi''(x^*))^2 g(x^*)} \int (K^{(1)}(\psi))^2 d\psi\right), \quad \text{(A.1-2)}
\]

where \( \mathcal{N}(0, \sigma^2) \) denotes a centered normal distribution with variance \( \sigma^2 \).

Let \( \{x_n\} \) be any sequence of reals in \( \mathcal{X} \). We next establish that \( \hat{\pi}''(x_n) \) converges to \( \pi''(x) \) whenever \( |x_n - x| \) converges to zero in probability.

\[
|\hat{\pi}''(x_n) - \pi''(x)| \leq |\hat{\pi}''(x_n) - \pi''(x_n)| + |\pi''(x_n) - \pi''(x)|
\]

\[
\leq \sup_{r_0 \leq r \leq r_{\max}} |\hat{\pi}''(x) - \pi''(x)| + |\pi''(x_n) - \pi''(x)|
\]

The second term on the right-hand-side converges to zero in probability by continuity of \( \pi''(\cdot) \). The first term on the right-hand-side also converges to zero under the current assumption that \( nh_n^6 \rightarrow \infty \) (see Ziegler (2002, Theorem 1.5)).

Now note that the sequence \( x_{1,n} \) converges in probability to \( x^* \) (since \( \hat{x}_n \) converges in probability to \( x^* \) by (A.1-2) and \( |x_{1,n} - x^*| \leq |\hat{x}_n - x^*| \)). Applying the
result above, we have that \( \hat{\pi}''(x_{1,n}) \) converges to \( \pi''(x^*) \) in probability. Noting that (A.1-1) implies that

\[
nh_n^3 A_n = -\frac{1}{2} \hat{\pi}''(x_{1,n}) \left[ \sqrt{nh_n^3} (\hat{x}_n - x^*) \right]^2 ,
\]

Slutsky's theorem and the continuous mapping theorem, in conjunction with (A.1-2), imply that

\[
nh_n^3 A_n \Rightarrow \left( -\frac{1}{2\pi''(x^*)} \frac{\sigma^2(x^*)}{g(x^*)} \int (K^{(1)}(\psi))^2 d\psi \right) \chi^2_1
\]

We now turn to analyze the second contribution to \( \Delta_n, B_n \). We start with a result on the estimate \( \hat{\theta} \). The next lemma provides a characterization \( \hat{\theta} \)'s convergence to \( \theta^* \) under both \( H_0 \) and \( H_1 \).

**Lemma 2** i.) Under \( H_0 \),

\[
\sqrt{n}(\hat{\theta} - \theta^*) \Rightarrow \mathcal{N}(0, \Sigma^2) \quad as \ n \to \infty, \tag{A.1-3}
\]

where \( \Sigma^2 = I_{\theta^*}^{-1} \) and \( I_{\theta^*} \) is the Fisher information matrix.

ii.) Under \( H_1 \),

\[
\hat{\theta} \to \theta^* \quad \text{almost surely as } n \to \infty, \tag{A.1-4}
\]

**Proof of Lemma 2.** For part i.), note that the conditions spelled out in Assumption 4 ensure asymptotic normality of the ML estimator (see, e.g., Serfling (2002)). For part ii.), Assumption 4 ensures that one can apply White (1982, Theorem 2.2) which yields the result.
We now characterize the asymptotic behavior of $B_n$ under the null and alternative hypotheses.

**Lemma 3**

i.) Under $H_0$, $nh_n^3 B_n \to 0$ in probability as $n \to \infty$.

ii.) Under $H_1$, $nh_n^3 B_n \to \infty$ in probability as $n \to \infty$.

**Proof of Lemma 3.** To prove the result, we decompose $B_n$ into two terms. We have $B_n = C_n + D_n$, where

$$C_n = \hat{\pi}_n(x^*) - \hat{\pi}_n(x^*(\theta^*))$$

$$D_n = \hat{\pi}_n(x^*(\theta^*)) - \hat{\pi}_n(x^*(\hat{\theta}))$$

**Analysis under $H_0$.** By a Taylor expansion, we have

$$-D_n = \hat{\pi}'_n(x^*(\theta^*))(x^*(\hat{\theta}) - x^*(\theta^*)) + \frac{1}{2} \hat{\pi}''_n(x_{1,n})(x^*(\theta^*) - x^*(\hat{\theta}))^2$$

for some $x_{1,n} \in [\min\{x^*(\theta^*), x^*(\hat{\theta})\}, \max\{x^*(\theta^*), x^*(\hat{\theta})\}]$.

We next establish that $x^*(\cdot)$ is Lipschitz continuous. Let $C_2 = \max\{|\partial^2 p(x^*(\theta); \theta)/\partial \theta_i \partial x| : i = 1, \ldots, d, \theta \in \Theta\}$ and $C_3 = \min\{|\partial p(x^*(\theta); \theta)/\partial^2 x| : i = 1, \ldots, d, \theta \in \Theta\}$. Note that $C_2$ and $C_3$ are well defined as the maximum and minimum of continuous functions over compact sets. In addition, note that by Assumption 3ii.) that $C_3 > 0$. The fact that $x^*(\theta)$ is an interior maximizer implies that it satisfies

$$\frac{\partial p(x^*(\theta); \theta)}{\partial x} = 0.$$
In addition, the fact that \( \frac{\partial^2 p(x^*(\theta); \theta)}{\partial x^2} < 0 \) (Assumption 3ii.) implies that the equation above has a unique solution in the neighborhood of \( x^*(\theta) \). Applying the implicit function theorem yields that \( x^*(\theta) \) is differentiable and

\[
\frac{\partial x^*(\theta)}{\partial \theta_i} = -\frac{\frac{\partial^2 p(x^*(\theta); \theta)}{\partial \theta_i \partial x}}{\frac{\partial^2 p(x^*(\theta); \theta)}{\partial^2 x} x'}. 
\]

We then have that \( |\partial x^*(\theta)/\partial \theta_i| \leq C_2/C_3 \) and \( x^*(\theta) \) is Lipschitz continuous with constant \( C_2/C_3 \).

Under \( H_0 \), \( x^*(\theta^*) = x^* \) and hence \( \pi'(x^*(\theta^*)) = 0 \). Coming back to (A.1-5), it follows that

\[
nh_n^3 D_n = -nh_n^3 \pi''(x^*(\theta^*)) (x^*(\theta) - x^*(\theta^*)) - nh_n^3 \frac{1}{2} \pi''(x^*(\theta^*)) (x^*(\theta) - x^*(\theta^*))^2 \\
= -h_n^{1/2} \sqrt{nh_n^3} \left[ \pi'(x^*(\theta^*)) - \pi'(x^*(\theta^*)) \right] \frac{\sqrt{n}}{h_n} (x^*(\theta) - x^*(\theta^*)) \\
- \frac{1}{2} \pi''(x^*(\theta^*)) nh_n^3 (x^*(\theta^*) - x^*(\hat{\theta}))^2. \tag{A.1-6}
\]

Focusing on the second on the right-hand-side above, one can establish as in Lemma 1 that \( \hat{\pi}''(x_{1,n}) \) converges to \( \pi''(x^*(\theta^*)) \). We also have that \( nh_n^3 (x^*(\theta^*) - x^*(\hat{\theta}))^2 \leq h_n^3 (C_2/C_3)^2 [\sqrt{n}(\hat{\theta} - \theta^*)]^2 \) and the right hand side converges to zero in probability by (A.1-3) and the fact that \( h_n \rightarrow 0 \).

Turning to the first term on the right-hand-side in (A.1-6), we have that

\[
h_n^{1/2} \sqrt{nh_n^3} (\pi'(x^*(\theta^*)) - \pi'(x^*(\theta^*)) \) converges to zero under the assumption that
\( nh_n^3 \rightarrow 0 \) (see Pagan and Ullah (1999, Theorem 4.3)). On another hand, \( h_n \sqrt{n} (x^*(\theta^*) - x^*(\hat{\theta})) \leq h_n (C_2/C_3) \sqrt{n}(\hat{\theta} - \theta^*) \) and again the right hand side converges to zero in probability by (A.1-3) and the fact that \( h_n \rightarrow 0 \). We deduce that

\[
nh_n^3 D_n \Rightarrow 0 \tag{A.1-7}
\]
Under $H_0$, $x^*(\theta^*) = x^*$, which implies that $C_n = 0$ and hence i.) follows.

**Analysis under $H_1$.** We now analyze $D_n$ and $C_n$ under $H_1$. Under $H_1$, it is clear that $D_n$ converges to zero in probability from (A.1-5), the continuity of $x^*(\theta)$ and the consistency (A.1-4). On another hand, we have

$$C_n = \hat{\pi}_n(x^*) - \hat{\pi}_n(x^*(\theta^*))$$

$$= \hat{\pi}_n(x^*) - \pi(x^*) + \pi(x^*) - \pi(x^*(\theta^*)) + \pi(x^*(\theta^*)) - \pi(x^*(\theta^*)) - \hat{\pi}_n(x^*(\theta^*)),$$

Both terms $\hat{\pi}_n(x^*) - \pi(x^*)$ and $\pi(x^*(\theta^*)) + \pi(x^*(\theta^*)) - \hat{\pi}_n(x^*(\theta^*))$ converge to zero by the consistency of the nonparametric estimator. We deduce that $C_n$ converges to $\pi(x^*) - \pi(x^*(\theta^*)) > 0$ in probability. Hence $B_n$ converges to $\pi(x^*) - \pi(x^*(\theta^*))$ and $nh^3 B_n \Rightarrow \infty$ since $nh^3 \rightarrow \infty$. ii) is now established and the proof of Lemma 3 is complete. ■

Combing the results of Lemmas 1 and 3, the result of the theorem follows by an application of Slutsky's theorem. This completes the proof.
Appendix B

Proofs for Chapter 3

B.1 Proofs of Main Results

Preliminaries and notation. For any real number $x$, $x^+$ will denote $\max\{x, 0\}$. $C_1, C_2, \ldots$ will be used to denote positive constants which are independent of a given demand function, but may depend on the parameters of the class $\mathcal{L}$ of admissible demand functions and on $x$ and $T$. A sequence $\{a_n\}$ of real numbers is said to increase to infinity at a polynomial rate if there exist a constant $\beta > 0$ such that $\lim \inf_{n \to \infty} a_n / n^\beta > 0$. To lighten the notation, we will occasionally omit the arguments of $J^s(x, T; \lambda)$ and $J^D_n(x, T|\lambda)$. In what follows we will make use of the following facts.

Fact 1. Recall the definition of $J^D(x, T|\lambda)$, the optimal value of the deterministic relaxation (3.3.5). First note that $J^D_n = nJ^D$. We will also use the fact that

$$\inf_{\lambda \in \mathcal{L}} J^D_n(x, T|\lambda) \geq m^D,$$

where $m^D = mT' > 0$ and $T' = \min\{T, x/M\}$. Indeed, for any $\lambda \in \mathcal{L}$, there is a price $q \in [\underline{p}, \bar{p}]$ such that $r(q) \geq m$. Consider the policy that applies $q$ on $[0, T']$
and then applies $p_\infty$ up until $T$. This solution is feasible since $\lambda(q)T' \leq MT' \leq x$.

In addition the revenues generated from this policy are given by $mT'$.

We provide below a result that generalizes Gallego and van Ryzin (1994, Proposition 2) and that is used throughout the proofs of the main results. Its proof can be found in Appendix B.2.

**Lemma 4** The solution to problem (3.3.5) is given by $p(s) = p^D := \max\{p^u, p^c\}$ for $s \in [0, T']$ and $p(s) = p_\infty$ for $s > T'$, where $p^u = \arg\max_{p \in [p, \bar{p}]} \{r(\lambda(p))\}$, $p^c = \arg\min_{p \in [p, \bar{p}]} |\lambda(p) - x/T|$ and $T' = \min\{T, x/\lambda(p^D)\}$. In addition, for all $\lambda \in \mathcal{L}$, the optimal value of (3.3.5), $J^D(x, T; \lambda)$, serves as an upper bound on $J^\pi(x, T; \lambda)$ for all $\pi \in \mathcal{P}$.

Before proceeding, we state the following lemma which is needed throughout the analysis and whose proof is deferred to Appendix B.2.

**Lemma 5** Suppose that $\mu \in [0, M]$ and $r_n \geq n^\beta$ with $\beta > 0$.

If $\epsilon_n = 2\eta^{1/2} M^{1/2} (\log n)^{1/2} r_n^{-1/2}$, then for all $n \geq 1$

$$
\mathbb{P}\left( N(\mu r_n) - \mu r_n > r_n \epsilon_n\right) \leq \frac{C}{n^n},
$$

$$
\mathbb{P}\left( N(\mu r_n) - \mu r_n < -r_n \epsilon_n\right) \leq \frac{C}{n^n},
$$

for some suitably chosen constant $C > 0$.

This lemma bounds the deviations of a Poisson process from its mean. It will be used to control for the estimates of the demand function evaluated at given prices.
B.1.1 Proofs of the results in Section 3.4

Proof of Proposition 1. Fix $\lambda \in \mathcal{L}$, $\eta = 2$. We consider the sequence of policies $\pi_n := \pi(\tau_n, \kappa_n)$, $n = 1, 2, \ldots$ defined by means of Algorithm 1. The proof is organized in four steps. The first step develops an expression for a lower bound on the expected revenues achieved by the proposed policy. This lower bound highlights the terms that need to be analyzed. The second step provides probabilistic bounds on the estimate $\hat{\rho}$ of the price $p_D$ and the associated revenue rate. This is done by controlling the deviations of a Poisson process from its mean (see Lemma 6). The third step finalizes the analysis of the lower bound. For this step, two cases have to be considered separately depending on the starting level of inventory. The key issue here is to control for stochastic fluctuations of customer requests at the estimated price $\hat{\rho}$. In the last step, we conclude the proof by combining the results from the two cases and plugging in the tuning parameters.

Let $\tau_n$ be such that $\tau_n \to 0$ and $n\tau_n \to \infty$ at a polynomial rate as $n \to \infty$. Let $\kappa_n$ be a sequence of integers such that $\kappa_n \to \infty$ and $n\Delta_n := n\tau_n/\kappa_n \to \infty$ at a polynomial rate. We divide the interval $[\bar{p}, \bar{p}]$ into $\kappa_n$ equal length intervals and we let $P_n = \{p_i, i = 1, \ldots, \kappa_n\}$ be the left endpoints of these intervals. Now partition $[0, \tau_n]$ into $\kappa_n$ intervals of length $\Delta_n$ and apply the price $p_i$ on the $i^{th}$ interval. Define

$$\hat{\lambda}(p_i) = \frac{N(\sum_{j=1}^{i} n\lambda(p_j)\Delta_n) - N(\sum_{j=1}^{i-1} n\lambda(p_j)\Delta_n)}{n\Delta_n}, \quad i = 1, \ldots, \kappa_n, \quad (B.1-1)$$

where $N(\cdot)$ is a unit rate Poisson process. Thus $\hat{\lambda}(p_i)$ denotes the number of
product requests over successive intervals of length $\Delta_n$, normalized by $n\Delta_n$.

**Step 1.** Here, we derive a lower bound on the expected revenues under the policy $\pi_n$. Let $X_n^{(L)} = \sum_{i=1}^{\kappa_n} \lambda(p_i)n\Delta_n$, $X_n^{(P)} = \lambda(\hat{p})n(T - \tau_n)$ and put $Y_n = N(X_n^{(L)} + X_n^{(P)})$, $Y_n^{(L)} = N(X_n^{(L)})$ and $Y_n^{(P)} = Y_n - Y_n^{(L)}$. $Y_n^{(L)}$ represents the maximum number of requests during the learning phase if the system would not run out of resources, $Y_n^{(P)}$ the maximum number of requests during the pricing phase and $Y_n$ the maximum total number of requests throughout the sales horizon. Now note that one can lower bound the revenues achieved by $\pi_n$ by those accumulated during the pricing phase and during the latter, the maximum number of units that can be sold is exactly $\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}$. We deduce that

$$J_n^\pi \geq \mathbb{E}\left[\hat{p}\min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\}\right] \quad (B.1-2)$$

Next, we analyze the lower bound above by getting a handle on $\hat{p}$, the estimate of $p^D$ and the quantities $Y_n^{(P)}$ and $Y_n^{(L)}$.

**Step 2.** Here, we analyze the estimate $\hat{p}$ through the estimates $\hat{p}^u$ and $\hat{p}^c$. Recall the definition of

$$\hat{p}^u = \arg\max_{1 \leq i \leq \kappa_n} \{p_i: \hat{\lambda}(p_i)\}, \quad \hat{p}^c = \arg\min_{1 \leq i \leq \kappa_n} |\hat{\lambda}(p_i) - x/T|,$$

which are estimates of $p^u$ and $p^c$, respectively. We define the following quantity

$$u_n = (\log n)^{1/2} \max\left\{1/\kappa_n, 1/(n\Delta_n)^{1/2}\right\}, \quad (B.1-3)$$

that will be used throughout the analysis to quantify deviations of various quantities associated with the proposed policy from their full information counterparts.
We provide below a result that characterizes the revenue rate at $\hat{p} = \max\{\hat{p}^c, \hat{p}^u\}$ as well as $\hat{p}^c$. The proof of this lemma can be found in Appendix B.2.

**Lemma 6** For some positive constants $C_1, C_2, C_3$, for and all $n \geq 1$,

\[
\mathbb{P}\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_1u_n\} \leq \frac{C_3}{n^{\eta-1}}, \quad (B.1-4)
\]
\[
\mathbb{P}\{|\hat{p}^c - p^c| > C_2u_n\} \leq \frac{C_3}{n^{\eta-1}}. \quad (B.1-5)
\]

**Step 3.** In what follows we will separate two cases: $\lambda(\bar{p}) \leq x/T$ and $\lambda(\bar{p}) > x/T$. The latter case is one where the decision-maker would on average run out of inventory during the sales horizon and has to be analyzed separately.

**Case 1.** Suppose first that $\lambda(\bar{p}) \leq x/T$. Note that $\min\{Y^{(P)}_n, (nx - Y^{(L)}_n)^+\} = Y^{(P)}_n - (Y_n - nx)^+$. Hence, using (B.1-2) and the fact that $\hat{p} \leq \bar{p}$, we have

\[
J_n = \mathbb{E}[\hat{p}Y^{(P)}_n] - \bar{p}\mathbb{E}[(Y_n - nx)^+]
\]
\[
= \mathbb{E}[\mathbb{E}[\hat{p}N(\lambda(\hat{p})n(T - \tau_n)) \mid \hat{p}]] - \bar{p}\mathbb{E}[(Y_n - nx)^+]
\]
\[
= \mathbb{E}[r(\lambda(\hat{p}))]n(T - \tau_n) - \bar{p}\mathbb{E}[(Y_n - nx)^+]. \quad (B.1-6)
\]

Now,

\[
\mathbb{E}[r(\lambda(\hat{p}))]
\]
\[
r(\lambda(p^D))
\]
\[
+ \mathbb{E}\left[r(\lambda(\hat{p})) - r(\lambda(p^D)) \mid r(\lambda(p^D)) - r(\lambda(\hat{p})) \leq C_1u_n\right]\mathbb{P}\{r(\lambda(p^D)) - r(\lambda(\hat{p})) \leq C_1u_n\}
\]
\[
+ \mathbb{E}\left[r(\lambda(\hat{p})) - r(\lambda(p^D)) \mid r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_1u_n\right]\mathbb{P}\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_1u_n\}
\]
\[
r(\lambda(p^D)) - C_1u_n - \frac{C_4}{n^{\eta-1}}, \quad (B.1-7)
\]
where (a) follows from Lemma 6 and one can take $C_4 = 2\bar{p}MC_3$ (recall that $\lambda(\cdot)$ is bounded above by $M$ under Assumption 1 (i) and hence $r(\cdot)$ is bounded above by $\bar{p}M$.) We now turn to analyze the second term on the RHS of (B.1-6). The following result gives a probabilistic bound on the deviations of the random variable $Y_n$ from $nx$ and its proof is deferred to Appendix B.2.

**Lemma 7** For some positive constants $C_5, C_6$, for and all $n \geq 1$,

\[
\mathbb{P}(Y_n - nx > C_5nu_n) \leq \frac{C_6}{n^{\eta-1}} \tag{B.1-8}
\]

Using (B.1-8), we have

\[
\mathbb{E} \left[(Y_n - nx)^+\right] \leq C_5nu_n + \mathbb{E} \left[(Y_n - nx)^+ ; Y_n - nx \geq C_5nu_n\right] \\
\leq C_5nu_n + (nx + C_5nu_n + 1 + nM) \frac{C_6}{n^{\eta-1}} \\
\leq C_7nu_n, \tag{B.1-9}
\]

where $C_7$ is a suitably large positive constant, (a) follows from the fact that for a Poisson random variable $Z$ with mean $\mu$, $\mathbb{E}[Z | Z > a] \leq a + 1 + \mu$. We thus conclude from (B.1-7) and (B.1-9) that

\[
J_n^\pi \geq \left[r(\lambda(p^D)) - C_1u_n - \frac{C_3}{n^{\eta}}\right]n(T - \tau_n) - \bar{p}C_7nu_n \\
\geq nr(\lambda(p^D))T - C_8n(u_n + \tau_n), \tag{B.1-10}
\]

where $C_8$ is a suitably chosen constant. Now, from (B.1-10) and the definition of $m^D$ in Fact 1 in the preamble of the appendix, we have that

\[
\frac{J_n^\pi}{J_n^D} = \frac{J_n^\pi}{nJ^D} \geq 1 - \frac{C_8}{m^D}(u_n + \tau_n). \tag{B.1-11}
\]
Case 2. Now suppose $\lambda(\bar{p}) > x/T$. In this case, note that $p^c = p^D = \bar{p}$.

Recalling the lower bound expression in (B.1-2), we seek to get a handle on
\[ \min\left\{ Y_n^{(P)}, (nx - Y_n^{(L)})^+ \right\}. \]

The following result, whose proof can be found in Appendix B.2, says that the previous quantity will be “close” to the total units in
inventory, $nx$, with very high probability.

**Lemma 8** Define $\mathcal{A} := \{ \omega : \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \geq nx - C_9 u_n, \ |\hat{p} - p^D| \leq C_2 u_n \}$. Then for some positive constants $C_9, C_{10}$, for and all $n \geq 1$,
\[ \mathbb{P}(\mathcal{A}) \geq 1 - \frac{C_{10}}{n^{n-1}} \]  
(B.1-12)

The revenues generated by $\pi_n$ can be bounded below as follows

\[ J_n^\pi \geq \mathbb{E}\left[ \hat{p} \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \bigg| \mathcal{A} \right] \mathbb{P}(\mathcal{A}) \]
\[ \geq \mathbb{E}\left[ (p^D - C_2 u_n) \min\{Y_n^{(P)}, (nx - Y_n^{(L)})^+\} \bigg| \mathcal{A} \right] \mathbb{P}(\mathcal{A}) \]
\[ \geq (p^D - C_2 u_n)(nx - C_9 u_n) \left( 1 - \frac{C_{10}}{n^{n-1}} \right) \]  
(B.1-13)
\[ \geq p^D nx - n C_{11} u_n \]  
(B.1-14)

where both (a) and (b) follow from the definition of $\mathcal{A}$ and (B.1-12) and where $C_{10} > 0$ is suitably large. Now, recall that since $\lambda(\bar{p}) > x/T$, $p^D = \bar{p}$ and $J_n^D = nx\bar{p}$.

Hence
\[ \frac{J_n^\pi}{J_n^D} = \frac{J_n^\pi}{nx\bar{p}} \geq 1 - \frac{C_{11}}{\bar{p} x} u_n. \]  
(B.1-15)

**Step 4.** Let $C_{12} = \max\{C_9/(m^D), C_{11}/(\bar{p} x)\}$. Combining (B.1-11) and (B.1-15), we have for all $\lambda \in \mathcal{L}$, $\mathcal{R}_n^\pi(x, T; \lambda) \leq C_{12}(u_n + \tau_n)$, and note that $C_{12}$ does not de-
pend on the specific function $\lambda \in \mathcal{L}$. Note that the choice of tuning parameters that minimizes the order of the the upper bound on the regret is exactly $\tau_n \approx n^{-1/4}$ and $\kappa_n \approx n^{1/4}$. Plugging in, we get for some $C_{13} > 0$

$$\sup_{\lambda \in \mathcal{L}} R_n^\tau(x, T; \lambda) \leq \frac{C_{13} (\log n)^{1/2}}{n^{1/4}}.$$ 

The proof is complete. ■

**Proof of Proposition 2.** Since $J_n^\tau(x, T; \lambda) \leq J_n^*(x, T; \lambda)$, the result will be established if we can prove that for some $C > 0$

$$\sup_{\lambda \in \mathcal{L}} \left( 1 - \frac{J_n^*(x, T; \lambda)}{J_n^D(x, T; \lambda)} \right) > \frac{C}{n^{1/2}}. \quad (B.1-16)$$

Consider the demand function $\lambda(p) = (a - bp)^+ + b \in [K, \overline{K}]$ and $a = \max\{2\overline{K}p, \overline{K}p + x/T\}$. Note in particular that $\lambda \in \mathcal{L}$ since $a \leq M$. In addition, for this demand function we have $p^u = p^c = \overline{p}$. By Proposition 1 in Gallego and van Ryzin (1994), any dynamic pricing policy will never price below $p^u$, hence the optimal dynamic pricing policy is to price at $p^u$ until $T$ or the time when inventory is depleted. We deduce that

$$J_n^*(x, T|\lambda) = p^u \mathbb{E}[N(n\lambda(p^u)T)] - p^u \mathbb{E}[(N(n\lambda(p^u)T) - nx)^+]$$

$$= nr(p^u)T - p^u \mathbb{E}[(N(nx) - nx)^+]$$

$$= J_n^D(x, T|\lambda) - \overline{p} \mathbb{E}[(N(nx) - nx)^+]$$

$$= J_n^D(x, T|\lambda) - \overline{p} nx e^{-nx} (nx)^n x \overline{p} x (nx)!.$$ 

Hence,

$$\frac{J_n^D(x, T|\lambda) - J_n^*(x, T|\lambda)}{J_n^D(x, T|\lambda)} = \frac{1}{n \overline{p} \lambda(\overline{p}) T} \overline{p} nx e^{-nx} (nx)^n x \overline{p} x (nx)!.$$
Using Sterling’s approximation, \( n^{1/2}(1 - J_n^*(x, T|\lambda)/J_n^D(x, T|\lambda)) \rightarrow C_1 > 0 \) as \( n \rightarrow \infty \) and hence (B.1-16) is established. This completes the proof.

B.1.2 Proofs of the results in Section 3.5

In this section, we define for any \( \theta \in \Theta, p^D(\theta) := \max\{p^u(\theta), p^c(\theta)\} \), where \( p^u(\theta) = \arg\max\{p\lambda(p; \theta) : p \in [\underline{p}, \overline{p}]\}, p^c(\theta) = \arg\min\{\lambda(p; \theta) - x/T : p \in [\underline{p}, \overline{p}]\} \). In addition, we define \( D_\lambda(\theta) = \{l : l = \lambda(p; \theta) \text{ for some } p \in [\underline{p}, \overline{p}]\} \).

**Proof of Proposition 3.** In what follows, we let \( \theta^* \) denote the true parameter value. The proof follows the structure of the proof of Proposition 1. The first step provides a lower on the revenues achieved by the proposed policy. The second step bounds the difference between the estimated parameter vector and its true value and analyzes the revenues at \( p^u(\hat{\theta}), p^c(\hat{\theta}) \) and \( \hat{\theta} = \max\{p^u(\hat{\theta}), p^c(\hat{\theta})\} \). Step 3 finalizes the analysis of the lower bound. The proof concludes with the resulting upper bound obtained for the regret.

**Step 1.** Here, we derive a lower bound on the expected revenues under the policy \( \pi_n \). Let \( X_n^{(L)} = \sum_{i=1}^{k} \lambda(p_i)n\Delta_n, X_n^{(P)} = \lambda(\hat{p})n(T - \tau_n) \) and put \( Y_n = N(X_n^{(L)} + X_n^{(P)}), Y_{n}^{(L)} = N(X_n^{(L)}) \) and \( Y_{n}^{(P)} = Y_n - Y_{n}^{(L)} \). As in the proof of Proposition 1, we have

\[
J_n^\pi \geq \mathbb{E}\left[\hat{\theta} \min\{Y_{n}^{(P)}, (nx - Y_{n}^{(L)}))^+\}\right] \quad (B.1-17)
\]

Next, we analyze the lower bound above by getting a handle on \( \hat{\theta} \), the estimate of \( p^D \) and the quantities \( Y_{n}^{(P)} \) and \( Y_{n}^{(L)} \).

**Step 2.** Here, we analyze the estimate \( \hat{\theta} \).
We first analyze $p^u(\hat{\theta})$. Note that for any $p \in [\underline{p}, \overline{p}]$ and $\theta, \theta' \in \Theta$, we have

$$|r(\lambda(p; \theta) - r(\lambda(p; \theta'); \theta')| = |p\lambda(p; \theta) - p\lambda(p; \theta')|$$

$$\leq \overline{p}|\lambda(p; \theta) - \lambda(p; \theta')|$$

$$\leq \overline{p}\overline{K}_2\|\theta - \theta'\|_\infty,$$

where (a) follows from Assumption 6(ii.). Now

$$0 \leq r(\lambda(p^u(\theta^*); \theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*)$$

$$= \left[r(\lambda(p^u(\theta^*); \theta^*) - \lambda(p^u(\theta^*); \hat{\theta}); \hat{\theta})\right] + \left[r(\lambda(\hat{\theta}); \hat{\theta}) - r(\lambda(p^u(\hat{\theta}); \hat{\theta}); \hat{\theta})\right]$$

$$+ \left[r(\lambda(p^u(\hat{\theta}); \hat{\theta}) - r(\lambda(p^u(\theta^*); \theta^*); \theta^*)\right]$$

$$\leq \overline{K}_2\|\theta - \theta^*\|_\infty + 0 + \overline{K}_2\|\hat{\theta} - \theta^*\|_\infty,$$

where (a) follows from (B.1-18) and the fact that $p^u(\hat{\theta})$ maximizes $r(\lambda(\cdot ; \hat{\theta}); \hat{\theta})$ on $[\underline{p}, \overline{p}]$. Hence, we have established that

$$0 \leq r(\lambda(p^u(\theta^*); \theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*) \leq 2\overline{K}_2\overline{p}\|\hat{\theta} - \theta^*\|_\infty. \quad \text{(B.1-19)}$$

We now turn to analyze $p^c(\hat{\theta})$. Note that under Assumption 5(i.),(ii.), $r(\cdot, \theta)$ is Lipschitz with constant $\overline{p} + MK^{-1}$. Using this,

$$|r(\lambda(p^c(\theta^*); \theta^*); \theta^*) - r(\lambda(p^c(\hat{\theta}); \theta^*); \theta^*)|$$

$$\leq (\overline{p} + MK^{-1})|\lambda(p^c(\theta^*); \theta^*) - \lambda(p^c(\hat{\theta}); \theta^*)|. \quad \text{(B.1-20)}$$

Now note that

$$|\lambda(p^c(\theta^*); \theta^*) - \lambda(p^c(\hat{\theta}); \theta^*)|$$

$$\leq |\lambda(p^c(\theta^*); \theta^*) - \lambda(p^c(\hat{\theta}); \hat{\theta})| + |\lambda(p^c(\theta^*); \hat{\theta}) - \lambda(p^c(\hat{\theta}); \theta^*)|$$

$$\leq |\lambda(p^c(\theta^*); \theta^*) - \lambda(p^c(\hat{\theta}); \hat{\theta})| + \overline{K}_2\|\hat{\theta} - \theta^*\|_\infty,$$
where (a) follows from Assumption 6(ii.). Let $h_2(l; \theta)$ denote the projection on $\mathcal{D}_\lambda(\theta)$, i.e., $h_2(l; \theta) = \max\{\min\{y, \lambda(p; \theta)\}, \lambda(\rho; \theta)\}$ and note that $\lambda(p^C(\theta); \theta) = h_2(x/T; \theta)$. It is easy to show that for all $y$, $h_2(y, \cdot)$ is Lipschitz with constant $K_2$ where the latter was defined in Assumption 6(ii.). Using this, we have $|\lambda(p^C(\theta^*); \theta^*) - \lambda(p^C(\hat{\theta}); \theta^*)| = |h_2(x/T; \theta^*) - h_2(x/T; \hat{\theta})|$, and we deduce that

$$|\lambda(p^C(\theta^*); \theta^*) - \lambda(p^C(\hat{\theta}); \theta^*)| \leq 2K_2\|\theta - \theta^*\|_\infty.$$  

Combining (B.1-20) and (B.1-21), we get

$$|r(\lambda(p^C(\theta^*); \theta^*) - r(\lambda(p^C(\hat{\theta}); \theta^*); \theta^*)| \leq 2(p + MK^{-1})K_2\|\theta - \theta^*\|_\infty.$$

We now turn to analyze the revenues achieved by $\hat{p} = p^D(\hat{\theta})$. For this purpose, we divide the analysis into two cases:

**Case 1.** Suppose first that $p^u(\theta^*) \geq p^C(\theta^*)$, i.e., $p^D(\theta^*) = p^u(\theta^*)$.

i.) If $p^u(\hat{\theta}) \geq p^C(\hat{\theta})$, then

$$|r(\lambda(p^D(\theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*)| = |r(\lambda(p^u(\theta^*); \theta^*) - r(\lambda(p^C(\hat{\theta}); \theta^*); \theta^*)|

\leq 2K_2\rho\|\theta - \theta^*\|_\infty.$$

ii.) If $p^u(\hat{\theta}) < p^C(\hat{\theta}) \leq p^u(\theta^*)$, then the fact that $r(\lambda(\cdot; \theta^*), \theta^*)$ is nondecreasing to the left of $p^u(\theta^*)$ implies that

$$|r(\lambda(p^D(\theta^*); \theta^*) - r(\lambda(p^C(\hat{\theta}); \theta^*); \theta^*)| = |r(\lambda(p^u(\theta^*); \theta^*) - r(\lambda(p^C(\hat{\theta}); \theta^*); \theta^*)|

\leq |r(\lambda(p^u(\theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*)|

\leq 2K_2\rho\|\theta - \theta^*\|_\infty.$$
iii.) If \( p^u(\hat{\theta}) < p^c(\hat{\theta}) \) and \( p^c(\hat{\theta}) > p^u(\theta^*) \), then

\[
|r(\lambda(p^D(\theta^*); \theta^*); \theta^*) - r(\lambda(\hat{p}; \theta^*); \theta^*)| = |r(\lambda(p^d(\theta^*); \theta^*); \theta^*) - r(\lambda(p^c(\hat{\theta}); \theta^*); \theta^*)| \\
\leq (\bar{p} + MK^{-1})|\lambda(p^u(\theta^*); \theta^*) - \lambda(p^c(\hat{\theta}); \theta^*)| \\
\leq \left( a \right) (\bar{p} + MK^{-1})|\lambda(p^c(\theta^*); \theta^*) - \lambda(p^c(\hat{\theta}); \theta^*)| \\
\leq \left( b \right) 2(\bar{p} + MK^{-1})K_2\|\bar{\theta} - \theta^*\|_\infty.
\]

where \( a \) follows from the fact that \( p^c(\theta^*) \leq p^u(\theta^*) < p^c(\hat{\theta}) \) and that \( \lambda(\cdot; \theta^*) \) is nonincreasing and \( b \) follows from (B.1-21).

**Case 2.** Suppose now that \( p^u(\theta^*) < p^c(\theta^*) \), implying that \( p^D(\theta^*) = p^f(\theta^*) \).

i.) If \( p^c(\hat{\theta}) \geq p^u(\hat{\theta}) \), then

\[
|r(\lambda(p^D(\theta^*); \theta^*); \theta^*) - r(\lambda(\hat{p}; \theta^*); \theta^*)| = |r(\lambda(p^c(\theta^*); \theta^*); \theta^*) - r(\lambda(p^c(\hat{\theta}); \theta^*); \theta^*)| \\
\leq 2(\bar{p} + MK^{-1})K_2\|\bar{\theta} - \theta^*\|_\infty.
\]

ii.) If \( p^c(\hat{\theta}) < p^u(\hat{\theta}) \), then

\[
r(\lambda(p^D(\theta^*); \theta^*); \theta^*) - r(\lambda(\hat{p}; \theta^*); \theta^*) = r(\lambda(p^c(\theta^*); \theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*) \\
= r(\lambda(p^c(\theta^*); \theta^*); \theta^*) - r(\lambda(p^c(\hat{\theta}); \theta^*); \theta^*) \\
+ r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*) \\
+ r(\lambda(p^c(\theta^*); \theta^*); \theta^*) - r(\lambda(p^u(\hat{\theta}); \theta^*); \theta^*) \\
\leq \left( a \right) 2(\bar{p} + MK^{-1})K_2 + 2K_2\bar{p}\|\bar{\theta} - \theta^*\|_\infty.
\]

where \( a \) follows from the fact that \( p^u(\theta^*) \) maximizes \( r(\lambda(\cdot; \theta^*); \theta^*) \) on \([\bar{p}, \bar{p}]\) and the inequalities (B.1-19) and (B.1-22). Combining the results from cases 1 and 2,
we have established that

\[
    r(\lambda(p^D(\theta^*); \theta^*); \theta^*) - r(\lambda(\hat{\theta}; \theta^*); \theta^*) \leq C_1 \|\hat{\theta} - \theta^*\|_{\infty}, \tag{B.1-23}
\]

where \( C_1 = (2(\bar{p} + MK^{-1})K_2 + 2K_2\bar{p}) \). We are now left with controlling \( \hat{\theta} \). This is done in Lemma 9, whose proof is deferred to Appendix B.2.

**Lemma 9** Under Assumption 6, for some \( C_2 > 0 \),

\[
    \mathbb{E}[\|\hat{\theta} - \theta^*\|_{\infty}] \leq \frac{C_2}{(n\tau_n)^{1/2}} \quad \text{for all } n \geq 1. \tag{B.1-24}
\]

Using the result above and (B.1-23), we have

\[
    \mathbb{E}[r(\lambda(\hat{\theta}; \theta^*) \mid \theta^*))] = r(\lambda(p^D(\theta^*); \theta^*); \theta^*) + \mathbb{E}[r(\lambda(\hat{\theta}; \theta^*); \theta^*) - r(\lambda(p^D(\theta^*); \theta^*); \theta^*)]
    \geq r(\lambda(p^D(\theta^*); \theta^*); \theta^*) - C_1C_2(n\tau_n)^{1/2}.
\]

**Step 3.** Finalizing the analysis of the lower bound can be conducted in a similar manner as in the proof of Proposition 1 (Step 3) by letting in this case \( u_n = (\log n)^{1/2}(n\tau_n)^{-1/2} \).

**Step 4.** As in the proof of Proposition 1 (Step 4), we get that for some \( C_3 > 0 \),

\[
    \left(1 - \frac{J_n}{J^D_n}\right) \leq \frac{C_3}{mD}\left(\tau_n + (\log n)^{1/2}(n\tau_n)^{-1/2}\right), \tag{B.1-25}
\]

and the result follows by plugging in \( \tau_n \propto n^{-1/3} \). This completes the proof.

**Proof of Proposition 5.** Throughout the proof, we let \( \theta^* \) denote the true underlying parameter value. Also, to simplify notation, we let \( p^D := p^D(\theta^*) \). The proof is organized as follows. We first lower bound the expected revenues achieved by the proposed policy. The key issue here is to account for the performance losses.
throughout the various stages of the algorithm. This is done by analyzing the estimates $\hat{\theta}^i$ and the prices $\hat{p}^i$ at each stage and controlling for the amount of inventory depleted as well as the revenues generated.

For $i = 1, ..., \ell$, let $X_n^i = \sum_{j=1}^{i} n\lambda(\hat{p}_j; \theta^*) \Delta_n^{(j)}$, and put $Y_n^i = N(X_n^i) - N(X_n^{i-1})$. Finally, let $Y_n = \sum_{i=1}^{\ell} Y_n^i$ and recall that $p^D(\theta) = \max(p^u(\theta), p^c(\theta))$. We restrict attention to the case $\lambda(\bar{p}; \theta^*) \leq x/T$ (the case of $\lambda(\bar{p}; \theta^*) > x/T$ can be handled by following arguments similar to those in the second case of Step 3 in the proof of Proposition 1). The total revenues over the selling horizon under the policy $\pi_n$ can be bounded below as follows

$$J_n^* \geq \mathbb{E}[\hat{p}_1 Y_n^1 + \sum_{i=2}^{\ell} \hat{p}_i Y_n^i] - \bar{p} \mathbb{E}[(Y_n - nx)^+]$$

$$= n\hat{p}_1 \lambda(\hat{p}_1; \theta^*) \Delta_n^{(1)} + \sum_{i=2}^{\ell} \mathbb{E} \left[ \hat{p}_i \left( n\lambda(\hat{p}_i; \theta^*) \Delta_n^{(i)} \right) \right] - \bar{p} \mathbb{E}[(Y_n - nx)^+]$$

$$= n r(\lambda(p^D; \theta^*); \theta^*) \Delta_n^{(1)} + \sum_{i=2}^{\ell} \mathbb{E} \left[ \hat{p}_i \lambda(\hat{p}_i; \theta^*) \right] \Delta_n^{(i)} - \bar{p} \mathbb{E}[(Y_n - nx)^+]$$

$$= n r(\lambda(p^D; \theta^*); \theta^*) \sum_{i=1}^{\ell_n} \Delta_n^{(i)} - n \left[ r(\lambda(p^D; \theta^*); \theta^*) - r(\lambda(\hat{p}_1; \theta^*); \theta^*) \right] \Delta_n^{(1)}$$

$$- n \sum_{i=2}^{\ell_n} \left( r(\lambda(p^D; \theta^*); \theta^*) - \mathbb{E}[r(\lambda(\hat{p}_i; \theta^*); \theta^*)] \right) \Delta_n^{(i)} - \bar{p} \mathbb{E}[(Y_n - nx)^+], \quad \text{1-26}$$

$$= r(\lambda(p^D; \theta^*); \theta^*) nT - n \left[ r(\lambda(p^D; \theta^*); \theta^*) - r(\lambda(\hat{p}_1; \theta^*); \theta^*) \right] \Delta_n^{(1)}$$

$$- n \sum_{i=2}^{\ell_n} \left( r(\lambda(p^D; \theta^*); \theta^*) - \mathbb{E}[r(\lambda(\hat{p}_i; \theta^*); \theta^*)] \right) \Delta_n^{(i)} - \bar{p} \mathbb{E}[(Y_n - nx)^+], \quad \text{1-27}$$

where $\hat{p}_i$, $i = 2, ..., \ell_n$ were defined in Algorithm 3 (Step 2(b)).

Note that $(r(\lambda(p^D; \theta^*); \theta^*) - r(\lambda(\hat{p}_1; \theta^*); \theta^*)) \Delta_n^{(1)} \leq r(\lambda(p^D; \theta^*); \theta^*) \Delta_n^{(1)} \leq C_1 n^\alpha - 1$, where $C_1/2 = M\bar{p}$ bounds the revenue rate. The main task is to derive bounds on $r(\lambda(p^D; \theta^*); \theta^*) - \mathbb{E}[r(\lambda(\hat{p}_i; \theta^*); \theta^*)]$. Using Lemma 9 and a parallel reasoning to
the one in Step 2 of the proof of Proposition 3, we have that for some $C_2 > 0$

$$r(\lambda(p^D; \theta^*); \theta^*) - \mathbb{E}[r(\lambda(\hat{p}; \theta^*); \theta^*)] \leq \frac{C_2}{(n\Delta_n^{(i-1)})^{1/2}},$$

and hence,

$$\sum_{i=2}^{\ell_n} \left( r(\lambda(p^D; \theta^*); \theta^*) - \mathbb{E}[r(\lambda(\hat{p}; \theta^*); \theta^*)] \right) \Delta_n^{(i)} \leq C_1 \sum_{i=2}^{\ell_n} \frac{\Delta_n^{(i)}}{(n\Delta_n^{(i-1)})^{1/2}}$$

$$= C_2 \beta^{1/2} \sum_{i=2}^{\ell_n} n^{a^{i-1}} \leq C_2 T^{1/2} \ell_n n^{a_{\ell_n-1}},$$

where the last inequality follows since $\Delta_n^{(\ell)} = \beta$ and hence $\beta \leq T$. We deduce that

$$n \left[ r(\lambda(p^D; \theta^*); \theta^*) - r(\lambda(\hat{p}; \theta^*); \theta^*) \right] \Delta_n^{(1)}$$

$$+ n \sum_{i=2}^{\ell_n} \left( r(\lambda(p^D; \theta^*); \theta^*) - \mathbb{E}[r(\lambda(\hat{p}; \theta^*); \theta^*)] \right) \Delta_n^{(i)} + \mathbb{P}(Y_n - nx)^+ \right)$$

$$\leq n \max \{ C_2 T^{1/2}; C_1 \} \ell_n n^{a_{\ell_n-1}} + \mathbb{P}(Y_n - nx)^+. \quad \text{(B.1-28)}$$

Let $\epsilon_n$ be a sequence of positive numbers to be defined later. We have

$$\mathbb{E}[(Y_n - nx)^+] = \mathbb{E}[(Y_n - nx)^+; Y_n - nx \leq n\epsilon_n] + \mathbb{E}[(Y_n - nx)^+; Y_n - nx > n\epsilon_n]$$

$$\leq n\epsilon_n + \mathbb{E}[(Y_n - nx)^+; Y_n - nx > n\epsilon_n] \mathbb{P}(Y_n - nx > n\epsilon_n)$$

$$\leq n\epsilon_n + (n\epsilon_n + nM + 1) \mathbb{P}(Y_n - nx > n\epsilon_n),$$

where (a) follows from the fact that for a Poisson random variable $Z$ with mean $\mu$, $\mathbb{E}[Z | Z > a] \leq a + 1 + \mu$. Now note that $nx = n(x/T)T \geq n\lambda(p^c; \theta^*) \sum_{i=1}^{\ell_n} \Delta_n^{(i)}$. (Note that under the assumption that $\lambda(p^c; \theta^*) \leq x/T$ and since $\lambda(\cdot; \theta^*)$ is continuous and decreasing, either $\lambda(p^c; \theta^*) \geq x/T$ in which case $\lambda(p^c; \theta^*) = x/T$ or
\[ \lambda(p; \theta^*) < x/T \text{ in which case } \lambda(p^c; \theta^*) < x/T. \] Hence we have

\[ \mathbb{P}(Y_n - nx > \ell_n n \epsilon_n) \]

\[ \leq \mathbb{P}\left\{ N(n \lambda(\hat{p}_i; \theta^*) \Delta_{n}^{(i)}) > n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + n \epsilon_n \right\} \]

\[ + \sum_{i=2}^{\ell_n} \mathbb{P}\left\{ N(n \lambda(\hat{p}_i; \theta^*) \Delta_{n}^{(i)}) > n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + n \epsilon_n \right\}. \quad \text{(B.1-29)} \]

Choose \( \epsilon_n = C_3 (\log n)^{1/2} \Delta_{n}^{(1)} = C_3 (\log n)^{1/2} n^{\alpha_n - 1} \), where \( C_3 > 0 \) is suitably large.

Note that \( \Delta_{n}^{(1)}/\epsilon_n \to 0 \) as \( n \to \infty \), hence by Lemma 5 the first term on the RHS of (B.1-29) is bounded from above by \( C_4/n^\eta \) for some \( C_4 > 0 \) and some \( \eta \geq 2 \).

Consider any term in the sum which constitutes the second term on the RHS of (B.1-29). Using the same reasoning as the one leading to (B.2-10), we have for some \( C_5, C_6 > 0 \)

\[ \mathbb{P}\left\{ \lambda(\hat{p}_i; \theta^*) - \lambda(p^c; \theta^*) > \frac{C_5}{(\Delta_{n}^{(i-1)})^{1/2}} \right\} \leq \frac{C_6}{n^{\eta - 1}}. \quad \text{(B.1-30)} \]

Now, for \( i = 2, \ldots, \ell \)

\[ \mathbb{P}\left\{ N(n \lambda(\hat{p}_i; \theta^*) \Delta_{n}^{(i)}) > n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + n \epsilon_n \right\} \]

\[ \leq \mathbb{P}\left\{ N(n \lambda(\hat{p}_i; \theta^*) \Delta_{n}^{(i)}) > n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + n \epsilon_n, \lambda(\hat{p}_i; \theta^*) \leq \lambda(p^c; \theta^*) + \frac{C_6}{(n \Delta_{n}^{(i-1)})^{1/2}} \right\} \]

\[ \quad + \mathbb{P}\left\{ \lambda(\hat{p}_i; \theta^*) - \lambda(p^c; \theta^*) > \frac{C_6}{(n \Delta_{n}^{(i-1)})^{1/2}} \right\} \]

\[ \stackrel{(a)}{\leq} \mathbb{P}\left\{ N(n \lambda(p^c; \theta^*) + n C_6 \Delta_{n}^{(i)}/(\Delta_{n}^{(i-1)})^{1/2}) > n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + n \epsilon_n \right\} + \frac{C_6}{n^{\eta - 1}} \]

\[ \leq \mathbb{P}\left\{ N(n \lambda(p^c; \theta^*) + C_6 n^{\alpha_x}) - (n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + C_6 n^{\alpha_x}) > -n C_6 n^{\alpha_x - 1} + n \epsilon_n \right\} + \frac{C_6}{n^{\eta - 1}} \]

\[ \leq \mathbb{P}\left\{ N(n \lambda(p^c; \theta^*) + C_6 n^{\alpha_x}) - (n \lambda(p^c; \theta^*) \Delta_{n}^{(i)} + C_6 n^{\alpha_x}) > \frac{1}{2} n \epsilon_n \right\} + \frac{C_6}{n^{\eta - 1}}, \quad \text{(b)} \]

where (a) follows from (B.1-30) and (b) follows from the definition of the sequence \( \Delta_{n}^{(i)}, i = 1, \ldots, \ell \). Appealing to Lemma 5 yields that the first term on the RHS
above can be bounded by $C_7/n^n$ for some $C_7 > 0$. We conclude that for some $C_8 > 0$,

$$
\mathbb{E}
\left[
(Y_n - n x)^+ \right] \leq n C_8 (\log n)^{1/2} n^{a t n - 1},
$$

and this in turn, in conjunction with (B.1-27) and (B.1-28), implies that for some $C_9 > 0$

$$
\frac{J_n^x}{J_D^x} \geq 1 - \frac{C_9}{\text{mD}} \left[ \ell_n n^{a t n - 1} + (\log n)^{1/2} n^{a t n - 1} \right].
$$

By taking $a_t n = 2^{t_n - 1}/(2^{t_n} - 1)$ with $\ell_n = (\log 2)^{-1} \log \log n$, we get that $n^{a t n - 1} \leq \exp(1)n^{-1/2}$. Plugging in (B.1-31), we get the final assertion of the proposition. A similar result holds when $\lambda(\bar{p}, \theta^*) > x/T$ and the the proof is thus complete.

**B.2 Proofs of Auxiliary Results**

In what follows, $C'_i$, $i = 1, 2, \ldots$ will denote positive constants that depend only on $x, T$ and the class $\mathcal{L}$ (but not on a specific function $\lambda \in \mathcal{L}$).

**Proof of Lemma 4.** Fix $\lambda \in \mathcal{L}$ and let $\mathcal{P}_\lambda$ denote the class of policies that “know” the demand function prior the the start of the selling season, and whose corresponding price process is non anticipating and satisfies the admissibility conditions (3.3.2)-(3.3.3). Put

$$
J^* (x, T | \lambda) := \sup_{\pi \in \mathcal{P}_\lambda} \mathbb{E} \left[ \int_0^T p(s) dN^\pi(s) \right],
$$

Problem (B.2-1) would be the one that the decision maker would solve if s/he would know the demand function $\lambda$ prior to the start of the selling season. Clearly
\( J^\pi(x, T; \lambda) \leq J^*(x, T|\lambda) \) for all \( \pi \in \mathcal{P} \). Hence, to establish the bound of the lemma, it is sufficient to show that \( J^*(x, T|\lambda) \leq J^D(x, T|\lambda) \). Consider the optimization problem given in (3.3.5) and its equivalent formulation in terms of rates:

\[
J^D(x, T|\lambda) = \max \left\{ \int_0^T r(\lambda(s)) ds : \int_0^T \lambda(s) ds \leq x, \lambda(s) \in [\bar{l}, \bar{l}] \cup \{0\} \text{ for all } 0 \leq s \leq T \right\}.
\]

(B.2-2)

If \( \bar{l} = 0 \), then the set of feasible rates is convex and Gallego and van Ryzin (1994, Theorem 2) yields the result. Suppose now that \( \bar{l} > 0 \) and let

\[
\lambda^c = \arg \min_{l \in [\bar{l}, \bar{l}]} |l - x/T|, \quad \lambda^u = \arg \max_{l \in [\bar{l}, \bar{l}]} r(l), \quad \lambda^D = \min\{\lambda^u, \lambda^c\}.
\]

Claim 1: If \( x/T \geq \bar{l} \), then setting \( \lambda(s) = \lambda^D \) for \( 0 \leq s \leq T \) is an optimal solution to the deterministic problem (B.2-2). In addition the value of the deterministic problem serves as an upper bound on the original stochastic problem.

To verify the claim we consider two cases. Suppose first that \( x/T \in [\bar{l}, \bar{l}] \), in which case \( \lambda^c = x/T \). If \( \lambda^u \leq x/T \), then applying \( \lambda^u \) throughout is feasible and is hence optimal. In addition, the optimal value of the deterministic problem is given by \( r(\lambda^u)T \) which is an upper bound on the stochastic problem (since \( r(l) \leq r(\lambda^u) \) for all \( l \in [\bar{l}, \bar{l}] \)). If \( \lambda^u > x/T \), then \( \lambda^u > \bar{l} \) and hence \( r(\cdot) \) is necessarily non-decreasing on \([0, \lambda^u]\) (by the concavity assumption). Thus if we relax the domain of feasible rates to the convex set \([0, \bar{l}]\), applying \( \lambda^D \) throughout is an optimal solution (see Gallego and van Ryzin (1994)). This is a feasible solution for the constrained problem and is hence optimal for that problem as well. Let \( J^D[0, \bar{l}] \) be
the value of the deterministic problem with feasible rates in \([0, \bar{l}]\) and let \(J^*[0, \bar{l}]\) be the value of the stochastic problem. We clearly have that

\[ J^D[l, \bar{l}] \overset{(a)}{=} J^D[0, \bar{l}] \overset{(b)}{\geq} J^*[0, \bar{l}] \overset{(c)}{\geq} J^*[l, \bar{l}], \]

where: (a) follows from the result we just established; (b) follows from Gallego and van Ryzin (1994, Theorem 2); and (c) is a clear consequence of constraint relaxation.

Suppose now that \(x/T > \bar{l}\). In this case, \(\lambda^u \leq \bar{l} < x/T\). Applying \(\lambda^u\) throughout is feasible and hence optimal. The optimal value of the deterministic problem is hence given by \(r(\lambda^u)T\) which is an upper bound on the stochastic problem. The assertion of the claim follows.

**Claim 2:** If \(x/T < \bar{l}\), then \(\lambda(s) = \lambda^D = \bar{l}\) for \(s \in [0, T']\) and \(\lambda(s) = 0\) for \(s \in (T', T]\), where \(T' = x/\bar{l}\), is an optimal solution to the deterministic relaxation.

To verify the claim, note that all units are sold at the maximum allowable price with the proposed solution and hence the revenues achieved are \(x\bar{p}\), which is a clear upper bound on the performance of any solution of the deterministic or the stochastic problem. Thus the assertion of the claim follows.

Combining the two claims yields the result of the lemma in the case \(l > 0\). •

**Proof of Lemma 5.** If \(\mu = 0\), then result is immediate. Now suppose \(\mu \in (0, M]\).

Using a Chernoff bound, we have for any \(\theta > 0\)

\[ \mathbb{P}\left( N(\mu r_n) - \mu r_n > r_n\epsilon_n \right) \leq \exp\left\{ -\theta r_n(\mu + \epsilon_n) + (\exp\{\theta\} - 1)\mu r_n \right\}. \] (B.2-3)

The expression in each of the exponents is minimized by setting \(\theta = \log(1 + \epsilon_n/\mu)\).
Plugging this into (B.2-3) yields

\[
P(N(\mu r_n) - \mu r_n > r_n \epsilon_n) \leq \exp\left\{-\log\left(1 + \frac{\epsilon_n}{\mu}\right) (\mu + \epsilon_n) + \epsilon_n\right\}
\]

\[
\leq \exp\left\{r_n \left( -\log\left(1 + \frac{\epsilon_n}{M}\right) (M + \epsilon_n) + \epsilon_n\right)\right\}. \quad (B.2-4)
\]

For the last inequality, note that the derivative of the term in the exponent with respect to \( \mu \) is given by \( -\log(1 + \epsilon_n/\mu) + \epsilon_n/\mu \), which is always positive for \( \epsilon_n > 0 \).

Now, using a Taylor expansion we get that for some \( \xi \in [0, \epsilon_n] \),

\[
-\log\left(1 + \frac{\epsilon_n}{M}\right)(M + \epsilon_n) + \epsilon_n = -\frac{1}{2} \frac{\epsilon_n^2}{1 + \xi M}.
\]

If \( \epsilon_n \leq 1 \), we have

\[
-\frac{1}{2} \frac{\epsilon_n^2}{1 + \xi M} \leq -\frac{\epsilon_n^2}{4M}.
\]

If \( \epsilon_n > 1 \), then

\[
-\frac{1}{2} \frac{\epsilon_n^2}{1 + \xi M} \leq -\frac{1}{2M} \frac{\epsilon_n}{1 + \epsilon_n} \leq -\frac{1}{4M}.
\]

Letting \( C(\eta) = 2\eta^{3/2}M^{1/2} \) and substituting for \( \epsilon_n \), we get

\[
P(N(\mu r_n) - \mu r_n > r_n \epsilon_n) \leq \exp\left\{-r_n \min\left\{\frac{1}{4M}, \frac{C^2(\eta) \log n}{4Mr_n}\right\}\right\}
\]

\[
= \max\left\{\exp\left\{-\frac{r_n}{4M}\right\}, \frac{1}{n}\right\},
\]

Now it is easy to check that \( \exp\left\{-r_n/4M\right\} \leq (4M \eta/\beta)^{1/\beta} \exp\left\{-\eta/\beta\right\}1/n^\eta \). Hence the first result follows with \( C_1 = \max\{1, (4M \eta/\beta)^{1/\beta} \exp\{-\eta/\beta\}\} \). The other inequality is established using identical arguments. This completes the proof. ■

**Proof of Lemma 6.** The proof is divided into three main steps. The first step analyzes key properties of the estimate \( \hat{p} \) and establishes (B.1-5). The second step
focuses on $\tilde{p}$, and the third step uses the results from the previous two steps to derive properties of $\hat{p} = \max\{\hat{p}^u, \hat{p}^c\}$ and establishes (B.1-4).

**Step 1.** Here, we focus on $\hat{p}^c$ and show the following. Put $C'_1 = (M + 2\bar{K})C'_2$ where $C'_2 = \frac{1}{\eta M^{1/2}} \max\{8\eta^{1/2}M^{1/2}, 2\bar{K}(\bar{p} - \bar{p})\}$, then for some $C'_3 > 0$ and for all $n \geq 1$

$$
\mathbb{P}\{|r(\lambda(\hat{p}^c)) - r(\lambda(\tilde{p}^c))| > C'_1 u_n\} \leq \frac{C'_3}{n^{\eta-1}}, \quad (B.2-5)
$$

$$
\mathbb{P}\{|\hat{p}^c - p^c| > C'_2 u_n\} \leq \frac{C'_3}{n^{\eta-1}}. \quad (B.2-6)
$$

For each $n$, let $[p_i, p_{i+1}]$ be the interval that contains $p^c$. Here, we drop the dependence on $n$ to avoid cluttering the notation. Now

$$
|\lambda(\hat{p}^c) - \lambda(p^c)| \leq |\lambda(\hat{p}^c) - \hat{\lambda}(\hat{p}^c)| + |\hat{\lambda}(\hat{p}^c) - \lambda(p^c)|
$$

$$
\leq |\lambda(\hat{p}^c) - \hat{\lambda}(\hat{p}^c)| + |\hat{\lambda}(p_i) - \lambda(p^c)| \quad (a)
$$

$$
\leq 2 \max_{1 \leq k \leq \kappa_n} |\hat{\lambda}(p_k) - \lambda(p_k)| + |\lambda(p_i) - \lambda(p^c)| \quad (b)
$$

$$
\leq 2 \max_{1 \leq k \leq \kappa_n} |\hat{\lambda}(p_k) - \lambda(p_k)| + \frac{\bar{K}(\bar{p} - \bar{p})}{\kappa_n} \quad (c)
$$

where $(a)$ follows from the definition of $\hat{p}^c = \arg\min_{1 \leq j \leq \kappa_n} |\hat{\lambda}(p_j) - \lambda(p^c)|$, $(b)$ follows from the triangle inequality, and $(c)$ follows from the fact that $\lambda(\cdot)$ is $\bar{K}$-
Lipschitz (Assumption 1(ii.)). We have

\[ \mathbb{P}\left\{ |\hat{p}^c - p^c| > C'_2 u_n \right\} \stackrel{(a)}{\leq} \mathbb{P}\left\{ |\lambda(\hat{p}^c) - \lambda(p^c)| > C'_2 K u_n \right\} \]

\[ \leq \mathbb{P}\left\{ \max_{1 \leq k \leq \kappa_n} |\lambda(p_k) - \lambda(p_k)| > \frac{(C'_2 K) u_n}{2} - \frac{K(\hat{p} - p)}{2\kappa_n} \right\} \]

\[ \leq \sum_{k=1}^{\kappa_n} \mathbb{P}\left\{ |\lambda(p_k) - \lambda(p_k)| > \frac{(C'_2 K) u_n}{4} \right\} \]

\[ \leq \frac{C'_3}{n^{\gamma-1}}, \]

where (a) follows from the fact that \( \gamma(\cdot) \) is \( K^{-1} \)-Lipschitz (Assumption 1(ii.)), (b) follows from a union bound and the fact that \( C'_2 \geq 2K(\hat{p} - p)K^{-1} \) and and (c) follows from Lemma 5 in conjunction with the inequality \( C'_2/(4K^{-1}) \geq 2\eta^{1/2}M^{1/2} \).

Note that (B.1-5) has now been established by setting \( C_3 = C'_2 \). The inequality for revenues now follows. Indeed, note that under Assumption 5, \( r(\lambda(\cdot)) \) is Lipschitz with constant \( M + \bar{p}K \). Hence, \(|r(\lambda(p^c)) - r(\lambda(\hat{p}^c))| \leq (M + \bar{p}K)|p^c - \hat{p}^c|\), and

\[ \mathbb{P}\left\{ |r(\lambda(p^c)) - r(\lambda(\hat{p}^c))| > C'_1 u_n \right\} \leq \mathbb{P}\left\{ |\hat{p}^c - p^c| > C'_1 u_n/(M + \bar{p}K) \right\} \]

\[ = \mathbb{P}\left\{ |\hat{p}^c - p^c| > C'_2 u_n \right\} \]

\[ \leq \frac{C'_3}{n^{\gamma-1}}. \]

This concludes the proof of (B.2-6) and (B.2-5). In Step 2, we derive properties of \( \hat{p}^u \) which are subsequently used in Step 3 to prove (B.1-4).

**Step 2.** Here, we focus on \( \hat{p}^u \) and show the following. Put

\[ C'_4 = \max\{8\eta^{1/2}M^{1/2}\bar{p}, 2M + \bar{p}K\}, \]

then for some \( C'_5 > 0 \) and all \( n \geq 1 \)

\[ \mathbb{P}\left\{ r(\lambda(p^u)) - r(\lambda(\hat{p}^u)) > C'_4 u_n \right\} \leq \frac{C'_5}{n^{\gamma-1}}, \]

(B.2-7)

First, for each \( n \), let \( j \) be the index such that \( p^u \in [p_j, p_{j+1}] \). Recall that under
Assumption 5, $r(\lambda(\cdot))$ is Lipschitz with constant $M + \bar{p}K$. This, in conjunction with the fact that $|p^n - p_{j+1}| \leq (\bar{p} - p)/\kappa_n$ yields

$$|r(\lambda(p^n)) - r(\lambda(p_{j+1}))| \leq \frac{(M + \bar{p}K)(\bar{p} - p)}{\kappa_n}. \quad (B.2-8)$$

We now establish an upper bound on the difference in revenues $r(\lambda(p^n)) - r(\lambda(\hat{p}^u))$ as follows

$$r(\lambda(p^n)) - r(\lambda(\hat{p}^u))$$

$$= r(\lambda(p^n)) - p_{j+1}\lambda(p_{j+1}) + p_{j+1}\lambda(p_{j+1}^n) - \hat{p}^u\lambda(\hat{p}^u) + p^n\lambda(\hat{p}^u) - \hat{p}^u\lambda(\hat{p}^u)$$

$$(a) \leq r(\lambda(p^n)) - p_{j+1}\lambda(p_{j+1}) + \hat{p}^u\lambda(\hat{p}^u) - \hat{p}^u\lambda(\hat{p}^u)$$

$$\leq |r(\lambda(p^n)) - r(\lambda(p_{j+1}))| + 2\max_{1 \leq k \leq \kappa_n} |p_k\lambda(p_k) - p_k\hat{\lambda}(p_k)|$$

$$(b) \leq \frac{(M + \bar{p}K)(\bar{p} - p)}{\kappa_n} + 2\max_{1 \leq k \leq \kappa_n} |p_k\lambda(p_k) - p_k\hat{\lambda}(p_k)|,$$

where $(a)$ follows from the definition of $\hat{p}^u$ given in (3.4.2), and $(b)$ follows from (B.2-8). Now

$$\mathbb{P}\left\{r(\lambda(p^n)) - r(\lambda(\hat{p}^u)) > C'_4u_n\right\}$$

$$\leq \mathbb{P}\left\{\max_{1 \leq k \leq \kappa_n} |p_k\lambda(p_k) - p_k\hat{\lambda}(p_k)| > \frac{u_n}{2} - \frac{(M + \bar{p}K)(\bar{p} - p)}{2\kappa_n}\right\}$$

$$(a) \leq \mathbb{P}\left\{\max_{1 \leq k \leq \kappa_n} |p_k\lambda(p_k) - p_k\hat{\lambda}(p_k)| > \frac{C'_4u_n}{4}\right\}$$

$$(b) \leq \sum_{k=1}^{\kappa_n} \mathbb{P}\left\{|\lambda(p_k) - \hat{\lambda}(p_k)| > \frac{C'_4u_n}{4\bar{p}}\right\}$$

$$(c) \leq \frac{\kappa_n C'_5}{n^\eta}$$

$$\leq \frac{C'_5}{n^\eta - 1},$$

where $(a)$ follows from the fact that $(C'_4/2)u_n \geq C'_3(\bar{p} - p)/\kappa_n$, $(b)$ follows from a
union bound and (c) follows from the fact that $C'_4/(4\bar{p}) \geq 2\eta^{1/2}M^{1/2}$ and Lemma 5. We have now established (B.2-7).

**Step 3.** In this last step, we now turn to analyze the revenue function evaluated at $\hat{p} = \max\{\hat{p}^u, \hat{p}^c\}$, which is an estimate of $p^D = \max\{p^u, p^c\}$. Define $C'_7 = \max\{2C'_1, 2C'_4, C'_2(M + \bar{K}\bar{p})\}$. We divide the analysis in two cases: $p^u \geq p^c$ and $p^u < p^c$.

**Case 1.** Suppose first that $p^u \geq p^c$, i.e., $p^D = p^u$.

i.) If $\hat{p}^u \geq \hat{p}^c$, then $r(\lambda(p^D)) - r(\lambda(\hat{p})) = r(\lambda(p^u)) - r(\lambda(\hat{p}^u))$.

ii.) If $\hat{p}^u < \hat{p}^c \leq p^u$, then $r(\lambda(p^D)) - r(\lambda(\hat{p})) = r(\lambda(p^u)) - r(\lambda(\hat{p}^c)) \leq r(\lambda(p^u)) - r(\lambda(\hat{p}^u))$, where the last inequality follows from the fact that $r(\lambda(\cdot))$ is non-decreasing on $[\hat{p}^u, p^u]$. Since, $C'_7 \geq C'_4$, in both i.) and ii.), we have by (B.2-7) that

$$\mathbb{P}\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C'_7 u_n\} \leq C'_5/n^{\eta - 1}.$$

iii.) The last subcase we analyze is when $\hat{p}^u < \hat{p}^c$ and $\hat{p}^c > p^u$. In this case, note that

$$\mathbb{P}\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C'_7 u_n\}
= \mathbb{P}\{r(\lambda(p^u)) - r(\lambda(\hat{p}^c)) > C'_4 u_n\}
\leq \mathbb{P}\{r(\lambda(p^u)) - r(\lambda(\hat{p}^c)) > C'_7 u_n ; |\hat{p}^c - p^c| \leq C'_2 u_n\}
+ \mathbb{P}\{|\hat{p}^c - p^c| > C'_2 u_n\}
\leq (a) \mathbb{P}\{M + K\bar{p})|p^u - \hat{p}^c| > C'_7 u_n ; |\hat{p}^c - p^c| \leq C'_2 u_n\} + \frac{C'_3}{n^{\eta - 1}}
\leq (b) \mathbb{P}\{M + K\bar{p})|p^c - \hat{p}^c| > C'_7 u_n ; |\hat{p}^c - p^c| \leq C'_2 u_n\} + \frac{C'_3}{n^{\eta - 1}}
\leq (c) \frac{C'_4}{n^{\eta - 1}},
where (a) follows the fact that $r(\lambda(\cdot))$ is $(M + \overline{Kp})$-Lipschitz under Assumption 1, and (B.2-6). Note that (b) stems from the fact that in case iii.), $|p^u - \hat{p}^c| \leq |p^c - \hat{p}^c|$ since $\hat{p}^c > p^u \geq p^c$ and (c) follows since $C_2'/(M + \overline{Kp}) \geq C_2'$. In all subcases, we have

\[ P\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_7'u_n\} \leq \frac{\max\{C_5', C_3'\}}{n^{\eta-1}}. \]

**Case 2.** We now turn to the case where $p^u < p^c$, implying that $p^D = p^c$.

i.) If $\hat{p}^c \geq \hat{p}^u$, then $r(\lambda(p^D)) - r(\lambda(\hat{p})) = r(\lambda(p^c)) - r(\lambda(\hat{p}^c))$. Hence,

\[ P\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_7'u_n\} = P\{r(\lambda(p^c)) - r(\lambda(\hat{p}^c))u_n > C_7'u_n\} \leq \frac{C_3'}{n^{\eta-1}}, \]

where (a) follows from (B.2-5) and the fact that $C_7' \geq C_1'$.

ii.) If $\hat{p}^c < \hat{p}^u$, then $r(\lambda(p^D)) - r(\lambda(\hat{p})) = r(\lambda(p^c)) - r(\lambda(\hat{p}^u)) = r(\lambda(p^c)) - r(\lambda(\hat{p}^c)) + r(\lambda(p^c)) - r(\lambda(p^u)) - r(\lambda(\hat{p}^u)) \leq r(\lambda(p^c)) - r(\lambda(\hat{p}^c)) + r(\lambda(p^u)) - r(\lambda(\hat{p}^u))$, where the inequality follows from the definition of $p^u$. We deduce that

\[ P\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_7'u_n\} \]

\[ \leq P\{r(\lambda(p^c)) - r(\lambda(\hat{p}^c)) + r(\lambda(p^u)) - r(\lambda(\hat{p}^u)) > C_7'u_n\} \]

\[ \leq P\{r(\lambda(p^c)) - r(\lambda(\hat{p}^c)) > C_7'u_n/2\} + P\{r(\lambda(p^u)) - r(\lambda(\hat{p}^u)) > C_7'u_n/2\} \]

\[ \leq \frac{C_5' + C_3'}{n^{\eta-1}}, \]

where (a) follows from (B.2-5) and (B.2-7) and the fact that $C_7' \geq 2\max\{C_1', C_4'\}$.

Putting together the results from cases 1 and 2, we have shown that

\[ P\{r(\lambda(p^D)) - r(\lambda(\hat{p})) > C_7'u_n\} \leq \frac{C_5' + C_3'}{n^{\eta-1}}, \]
which establishes (B.1-4). The proof is now complete. ■

**Proof of Lemma 7.** Put $C'_1 = 4 \max \{ \overline{K}T C_2, M, 2 \eta^{1/2} M^{1/2} \}$ where $C_2$ was defined in Lemma 6. First, using the fact that $Y_n = Y_n^{(L)} + Y_n^{(P)}$, note that

\[
P(Y_n - nx > C'_1 n u_n) \leq \mathbb{P}(Y_n^{(L)} > C'_1 n u_n/2) + \mathbb{P}(Y_n^{(P)} > nx + C'_1 n u_n/2). \tag{B.2-9}
\]

Prior to bounding the second term on the RHS above, note that since $p^c - \hat{p} = p^c - \max\{\hat{p}^u, \hat{p}^c\} \leq p^c - \hat{p}^c$ and hence

\[
P\{p^c - \hat{p} > C_2 u_n\} \leq P\{p^c - \hat{p}^c > C_2 u_n\} \leq \frac{C_3}{n^{\eta-1}},
\]

where the last inequality follows from Lemma 6. Let $C'_2 = \overline{K} C_2$. Recalling that $\lambda(\cdot)$ is monotone decreasing and $\overline{K}$-Lipschitz, we have that $\lambda(\hat{p}) - \lambda(p^c) > C'_2 u_n$ implies that $p^c - \hat{p} > C_2 u_n$. Hence

\[
P\{\lambda(\hat{p}) > \lambda(p^c) + C'_2 u_n\} \leq P\{p^c - \hat{p} > C_2 u_n\} \leq \frac{C_3}{n^{\eta-1}}. \tag{B.2-10}
\]

Coming back to the second term on the RHS of (B.2-9), we have

\[
P(Y_n^{(P)} > nx + C'_1 n u_n/2) \leq \mathbb{P}(N \left( (\lambda(\hat{p}) n(T - \tau_n) \right) > nx + C'_1 n u_n/2, \lambda(\hat{p}) \leq \lambda(p^c) + C'_2 u_n) + \mathbb{P}(\lambda(\hat{p}) > \lambda(p^c) + C'_2 u_n) \leq \frac{C_3}{n^{\eta-1}},
\]

\[
\overset{(a)}{\leq} \mathbb{P}\left( N \left( (\lambda(p^c) + C'_2 u_n) nT \right) > nx + C'_1 n u_n/2 \right) + \frac{C_3}{n^{\eta-1}},
\]

\[
\overset{(b)}{\leq} \mathbb{P}\left( N \left( (x/T + C'_2 u_n) nT \right) > nx + C'_1 n u_n/2 \right) + \frac{C_3}{n^{\eta-1}},
\]

\[
= \mathbb{P}\left( N \left( x(1 + C'_2 u_n/x) n \right) - x(1 + C'_2 u_n/ x) n > -C'_2 T u_n + C'_1 n u_n/2 \right) + \frac{C_3}{n^{\eta-1}},
\]
where (a) follows from (B.2-10), and (b) follows from the fact that \( \lambda(p^c) \leq x/T \). For the latter, note that under the assumption that \( \lambda(\bar{p}) \leq x/T \) and since \( \lambda(\cdot) \) is continuous and decreasing, either \( \lambda(p) \geq x/T \) in which case \( \lambda(p^c) = x/T \) or \( \lambda(p) < x/T \) in which case \( \lambda(p^c) < x/T \). Now, using the fact that \( C_1' - C_2'T \geq C_2'T \), we have

\[
\mathbb{P}\left( N\left( \lambda(p^c) + u_n\right)nT \right) > nx + C_1'n u_n/2 \right) 
\leq \mathbb{P}\left( N\left( x(1 + C_2'T u_n/x)n \right) - x(1 + C_2'T u_n/x)n > C_2'T u_n \right) 
\leq \frac{C_3'}{n^\eta}, \quad (B.2-11)
\]

where \( C_3' \) is suitably large and we have used Lemma 5 for the last inequality (since \( u_n/((\log n)^{1/2}n^{-1/2}) \) as \( n \to \infty \)). The first term on the RHS of (B.2-9) can be bounded using Lemma 5.

\[
\mathbb{P}\left( Y_n^{(L)} > C_1'n u_n/2 \right) = \mathbb{P}\left( N\left( \sum_{i=1}^{\kappa_n} \lambda(p_i)n \Delta_n \right) > C_1'n u_n/2 \right) 
\leq \sum_{i=1}^{\kappa_n} \mathbb{P}\left( N\left( \lambda(p_i)n \Delta_n \right) > \frac{C_1'n u_n}{2\kappa_n} \right) 
\leq (a) \kappa_n \mathbb{P}\left( N(Mn \Delta_n) - Mn \Delta_n > \frac{C_1'n u_n}{4\kappa_n} \right) 
\leq (b) \kappa_n \frac{C_3'}{n^\eta} \leq \frac{C_4'}{n^\eta-1},
\]

where \( C_4' > 0 \) is suitably large, (a) follows from the fact that
\( C_1'n u_n/(4\kappa_n) \geq Mn \Delta_n \), (b) follows from Lemma 5 and the fact that \( C_1'n u_n/(4\kappa_n) \geq 2\eta^{1/2}M^{1/2}(\log n)^{1/2}(n \Delta_n)^{1/2} \) and (c) follows from the fact that \( \kappa_n = o(n) \). Now, combining the above, we have

\[
\mathbb{P}\left( Y_n - nx > C_1'n u_n \right) \leq \frac{C_3 + C_3' + C_4'}{n^\eta-1}. \quad (B.2-12)
\]
The result is now established.

**Proof of Lemma 8.** Note that $A = \{ \omega : Y_n^{(P)} \geq nx - C_9nu_n, \ Y_n^{(L)} \leq C_9nu_n, \ |\hat{p} - p^D| \leq C_2u_n \}$. Now,

$$\mathbb{P}(A^c) \leq \mathbb{P}(Y_n^{(P)} < nx - C_9nu_n) + \mathbb{P}(Y_n^{(L)} > C_9nu_n) + \mathbb{P}(|\hat{p} - p^D| > C_2u_n) \leq \mathbb{P}(Y_n^{(P)} < nx - C_9nu_n) + \mathbb{P}(Y_n^{(L)} > C_9nu_n) + \frac{C_3}{n^{\eta - 1}}, \quad (B.2-13)$$

where (a) follows from a union bound and (b) follows from Lemma 6 (since $|\hat{p} - p^D| = |\hat{p} - \bar{p}| \leq |\hat{p} - p^D| = |\hat{p} - p^D|$). We now focus on the first term on the RHS of (B.2-13).

$$\mathbb{P}(Y_n^{(P)} < nx - C_9nu_n)$$

$$= \mathbb{P}(N(\lambda(\hat{p})n(T - \tau_n)) < nx - C_9nu_n)$$

$$\leq \mathbb{P}(N(\lambda(\bar{p})n(T - \tau_n)) < n\lambda(\bar{p})T - n\epsilon_n)$$

$$\leq \mathbb{P}(N(\lambda(\bar{p})n(T - \tau_n)) - \lambda(\bar{p})n(T - \tau_n) < n(M\tau_n - C_9u_n)) \leq \frac{C'_1}{n^{\eta - 1}},$$

where (a) follows from Lemma 5 as long as $C_9$ is chosen large enough and $C'_1 > 0$ is suitably large. We now turn to the second term on the RHS of (B.2-13).

$$\mathbb{P}(Y_n^{(L)} > C_9nu_n) \leq \sum_{i=1}^{\kappa_n} \mathbb{P}(\lambda_i n\Delta_n > C_9n/\kappa_nu_n)$$

$$\leq \sum_{i=1}^{\kappa_n} \mathbb{P}(\lambda_i n\Delta_n - \lambda(p_i)n\Delta_n > C_9n/\kappa_nu_n - Mn\Delta_n) \leq \frac{C'_2}{n^{\eta - 1}},$$

where (a) follows from Lemma 5 as long as $C_9$ is chosen large enough and $C'_2 > 0$
is suitably chosen. Coming back to (B.2-13), we get

\[ P(\mathcal{A}^c) \leq \frac{C_1' + C_2' + C_3'}{n^{\eta - 1}}, \]  

(B.2-14)

and the result follows. ■

**Proof of Lemma 9.** Consider each interval \([(i - 1)\Delta_n, i\Delta_n]\) for \(i = 1, ..., k\).

Suppose each is divided into \([n\Delta_n]\) intervals of length \(1/n\) each and measurements are taken every \(1/n\) units of time. Let \(X_n^i = (x_1^i, ..., x_{[n\Delta_n]}^i)\) denote the vector of total demand observed in each interval. Focusing on the \(i^{th}\) interval where price \(p_i\) is applied, and letting \(\mu_i = \lambda(p_i, \theta)\), the log-likelihood function can be written as

\[ L_n^i(X_n^i, \mu_i) = \sum_{j=1}^{[n\Delta_n]} \log f_{\mu_i}(x_j^i). \]

This expression is maximized by \(\hat{\mu}_i = \sum_{j=1}^{[n\Delta_n]} x_j^i/\Delta_n\). This follows since for a Poisson process with unknown constant intensity \(\mu\), given \(n\) observations of the number of events in non-overlapping intervals of unit length \((y_1, ..., y_n)\), the ML estimate \(\hat{\mu}\) is given by \(\hat{\mu} = (y_1 + ... + y_n)/n\).

Now note that for a Poisson process with unknown intensity \(\mu \in (l_0, M]\), the Fisher information \(I\) is bounded and positive, where \(I\) is given by

\[ I = \mathbb{E}_\mu \left[ \partial \log f_\mu(x)/\partial \mu \right]^2. \]

In addition, note that for any real number \(s \geq 2\), we have \(\eta := \mathbb{E}_\mu |\partial \log f_\mu(x)/\partial \mu|^s = \sup_{\mu \in (l_0, M]} \mu^{-s} \mathbb{E}_\mu |x - \mu|^s < \infty\). Hence, under Assumption 6, the conditions of Theorem 3 in Borovkov (1998, II.36) are satisfied and we have that there exists values \(0 < c < \infty\) and \(\delta > 0\) such that for all \(v\) and \(n \geq 1\),

\[ P\left(\left(\frac{n\Delta_n}{\mu_i - \mu_i^*}\right) \leq c e^{-\delta v^2}. \right) \leq ce^{-\delta v^2}. \]  

(B.2-15)
Setting $\mu_i = \lambda(p_i, \theta^*)$ where $\theta^*$ is the true parameter and using condition (i.)(b) in Assumption 6, we deduce that

$$
\mathbb{P}\left( (n\Delta_n)^{1/2} \|\hat{\theta} - \theta^*\|_\infty \geq \alpha v \right) \leq \sum_{i=1}^{k} \mathbb{P}\left( (n\Delta_n)^{1/2} |\hat{\mu}_i - \mu_i^*| \geq v \right) \leq cke^{-\delta v^2}, \quad (B.2-16)
$$

and in turn,

$$
\mathbb{E}\|\hat{\theta} - \theta^*\|_\infty = \int_0^\infty \mathbb{P}\left\{ (n\Delta_n)^{1/2} \|\hat{\theta} - \theta^*\|_\infty > (n\Delta_n)^{1/2}s \right\} ds
\leq a + \int_a^\infty cke^{-\delta n \Delta_n s^2/\alpha^2} ds
\leq a + \frac{ck\alpha^2}{2\delta n \Delta_n \alpha^2} e^{-\delta n \Delta_n \alpha^2/\alpha^2}.
$$

Taking $a = (n\Delta_n)^{-1/2}$, one gets

$$
\mathbb{E}\|\hat{\theta} - \theta^*\|_\infty \leq \frac{C'_1}{(n\tau_n)^{1/2}}, \quad (B.2-17)
$$

where $C'_1 = 1 + ck\alpha^2/(2\delta)e^{-\delta/\alpha^2}$. This concludes the proof. \qed
Appendix C
Proofs for Chapter 4

C.1 Proofs of Main Results

Notation. In what follows, if \( x \) and \( y \) are two vectors, \( x \preceq y \) if and only if \( x_i > y_i \) for at least one component \( i \); \( x^+ \) will denote the vector in which the \( i^{th} \) component is \( \max\{x_i, 0\} \). We define \( \bar{a} := \max\{a_{ij} : 1 \leq i \leq m, 1 \leq j \leq d\} \), where \( a_{ij} \) are the entries of the capacity consumption matrix \( A \). \( C_i, i \geq 1 \) will denote positive constants which are independent of a given demand function, but may depend on the parameters of the class of admissible demand functions \( \mathcal{L} \) and on \( A, x \) and \( T \). Recall that \( e \) denotes the vector of ones in \( \mathbb{R}^d \). For a sequence \( \{a_n\} \) of real numbers, we will say it converges to infinity at a polynomial rate if there exist \( \beta > 0 \) such that \( \liminf_{n \to \infty} a_n/n^\beta > 0 \). With some abuse of notation, for a vector, \( y \in \mathbb{R}_+^d \) and a \( d \)-vector of unit rate Poisson processes \( N(\cdot) \), we will use for \( N(y) \) to denote the vector with \( i^{th} \) component \( N^i(y^i) \), \( i = 1, \ldots, d \). Finally

Comment 1. Recall the definition of problem (4.2.4). Since \( D_p \) is bounded, the price charged for any product never exceeds, say \( \bar{M} \). Consider a system where backlogging is allowed in the following sense: for each unit of resource backlogged
the system incurs a penalty of $M$. Recall that $A$ is assumed to be integer valued with no zero column, and hence anytime the new system receives a request such that no sufficient resources are available to fulfill it, a penalty of at least $M$ is incurred. Consider any admissible policy $\pi$ that applies $p_\infty$ for the remaining time horizon as soon as one resource is out of stock. (Note that all the policies introduced in the main text are of this form.) Since $M$ exceeds the price that the system receives, the expected revenues of such a policy $\pi$ in the original system $J^\pi(x, T; \lambda)$ are bounded below by the ones in the new system (note that in the latter, $\pi$ does not apply $p_\infty$ if the system runs out of any resource).

Comment 2. We will denote by $J^D(x, T|\lambda)$ the optimal value of the deterministic relaxation (4.4.2). First note that $J^D_n = nJ^D$. We will also use the fact that

$$\inf_{\lambda \in \mathcal{L}} J^D(x, T|\lambda) \geq m^D,$$

where $m^D = mT' > 0$, and $T' = \min\{T, \min_{1 \leq i \leq \ell} x_i/(\bar{a}Md)\}$. Indeed, for any $\lambda \in \mathcal{L}$, there is a price $q \in \mathcal{D}_p$ such that $r(q) \geq m$. Consider the policy that applies $q$ on $[0, T')$ and then applies $p_\infty$ up until $T$. This solution is feasible since $A\lambda(q)T' \leq d\bar{a}MT'e \leq x$. In addition the revenues generated from the policy above are given by $mT'$.

We provide below a lemma that will be used in the upcoming proofs. Its proof can be found in Appendix C.2.

**Lemma 10** Fix $\eta > 0$. Suppose that $\mu_j \in (0, M)$, $j = 1, ..., d$. and $r_n = n^\beta$ with $\beta > 0$. Then, if $c_n = C(\eta)(\log n)^{1/2}r_n^{-1/2}$ with $C(\eta) = 2d\eta^{1/2}\bar{a}M^{1/2}$, then the
following holds
\[
\mathbb{P}\left(A(N(\mu r_n) - \mu r_n) \geq \tau_n \varepsilon_n e\right) \leq \frac{C_1}{n^\eta},
\]
\[
\mathbb{P}\left(A(N(\mu r_n) - \mu r_n) \geq -\tau_n \varepsilon_n e\right) \leq \frac{C_1}{n^\eta},
\]
where \(C_1 > 0\) is an appropriately chosen constant.

**Proof of Theorem 4.** Fix \(\lambda \in \mathcal{L}\) and \(\eta \geq 1\). Denote by \(\{\lambda_1, ..., \lambda_k\}\) the intensities corresponding to the prices \(\{p_1, ..., p_k\}\). Let \((P_0)\) denote the following linear optimization problem
\[
\max \left\{ \sum_{i=1}^{k} p_i \cdot \lambda_i t_i : \sum_{i=1}^{k} A \lambda_i t_i \leq x, \sum_{i=1}^{k} t_i \leq T, t_i \geq 0, i = 1, ..., k. \right\}
\]
The optimal value of \((P_0)\), \(V_{(P_0)}^*\), is known to be an upper bound to \(J^*\) (cf. Gallego and van Ryzin (1997, Theorem 1)). For a system with "size" \(n\), the optimal value is just \(n\) times the optimal value of the system with size 1, and the optimal solutions are the same. In what follows, for any feasible vector \(t\), we use \(V_{(P_0)}(t)\) to denote the value of the objective function.

**Step 1.** We first focus on the the learning and optimization phases. Let \(\tau_n\) be such that \(\tau_n = o(1)\) and \(n \tau_n \to \infty\) as \(n \to \infty\) at a polynomial rate. Divide \(\tau_n\) into \(k\) intervals of equal length \(\Delta_n = \tau_n/k\). Apply each feasible price during \(\Delta_n\) time units. Let
\[
\hat{\lambda}(p_i) = \frac{N \left(n \Delta_n \sum_{j=1}^{i} \lambda_j\right) - N \left(n \Delta_n \sum_{j=1}^{i-1} \lambda_j\right)}{n \Delta_n}, \quad i = 1, ..., k.
\]
Let \((\hat{P})\) denote the following linear optimization problem
\[
\max \left\{ \sum_{i=1}^{k} p_i \cdot \hat{\lambda}(p_i) t_i : \sum_{i=1}^{k} A \hat{\lambda}(p_i) t_i \leq x, \sum_{i=1}^{k} t_i \leq T - \tau_n, t_i \geq 0, i = 1, ..., k. \right\}.
\]
For $n$ sufficiently large, the feasible set of ($\hat{P}$) is nonempty (since $\tau_n = o(1)$) and compact and hence the latter admits an optimal solution, say $\hat{t}$. In what follows, for any feasible vector $t$, we use $V_{(\hat{P})}(t)$ to denote the value of the objective function.

**Step 2.** Here, we derive a lower bound on the expected revenues under the policy $\pi$. Consider applying the solution $\hat{t}$ to the stochastic system on the interval $(\tau_n, T]$. Let $\bar{M} := \max\{\|p_1\|, \ldots, \|p_k\|\}$ and define $X_n^{(L)} := \sum_{i=1}^{k} n\lambda(p_i)\Delta_n$, $X_n^{(i)} := \sum_{j=1}^{i} n\lambda_j\hat{t}_j$, $i = 1, \ldots, k$. Finally put $Y_n = AN(X_n^{(L)} + X_n^{(k)})$. As noted in the preamble of the appendix, one can lower bound $J_n^*$ as follows

$$J_n^* \geq \mathbb{E}\left[\sum_{i=1}^{k} p_i \cdot \left[N\left(X_n^{(L)} + X_n^{(i)}\right) - N\left(X_n^{(L)} + X_n^{(i-1)}\right)\right]\right] - \bar{M} \cdot \mathbb{E}\left[(Y_n - nx)^+\right],$$

where the equality follows from the fact that that given $\hat{t}$, $N\left(X_n^{(L)} + X_n^{(i)}\right) - N\left(X_n^{(L)} + X_n^{(i-1)}\right)$ is distributed as a Poisson random variable with mean $\lambda_i\hat{t}_i$.

Let $\delta_n := C_1 (\log n)^{1/2} (n\Delta_n)^{-1/2}$ with $C_1 > 0$ to be specified later and $\mathcal{H} := \left\{\omega : \max_{1 \leq i \leq k} \|\lambda_i - \lambda(p_i)\|T \leq \delta_n\right\}$. Since revenues are non-negative, we can lower bound the first sum in (C.1-1) above as follows

$$\sum_{i=1}^{k} p_i \cdot \lambda_i \mathbb{E}[\hat{t}_i] \geq \mathbb{E}\left[\sum_{i=1}^{k} p_i \cdot \lambda_i \hat{t}_i \left| \mathcal{H}\right\}\right] \mathbb{P}(\mathcal{H}).$$

**Lemma 11** For $\omega \in \mathcal{H}$, $\hat{t}$ is feasible for ($\hat{P}_0$) and for $C_2, C_3 > 0$ suitably large, we have

$$V_{(\hat{P}_0)}(\hat{t}) \geq V_{(\hat{P})}(\hat{t}) - C_2\delta_n, \quad \text{(C.1-2)}$$

$$V_{(\hat{P})}(\hat{t}) \geq V_{(\hat{P}_0)}^* - C_3 \max\{\delta_n, \tau_n\}. \quad \text{(C.1-3)}$$


We deduce that

\[
\mathbb{E}\left[ \sum_{i=1}^{k} p_i \cdot \lambda_i \hat{t}_i \mid \mathcal{H} \right] = \mathbb{E}\left[ V_{(P_0)}(\hat{t}) \mid \mathcal{H} \right] \geq (a) \mathbb{E}\left[ V_{(P)}(\hat{t}) - C_2 \delta_n \mid \mathcal{H} \right] \geq (b) V_{(P_0)} - (C_2 + C_3) \max\{\delta_n, \tau_n\},
\]

where (a) follows from (C.1-2) and (b) follows from (C.1-3). We now turn to bound the probability of the event \( \mathcal{H}^c \)

\[
\mathbb{P}(\mathcal{H}^c) \leq (a) \mathbb{P}(\max_{1 \leq i \leq k} \|\lambda_i - \hat{\lambda}(p_i)\|T > \delta_n) \leq (b) \sum_{i=1}^{k} \mathbb{P}(\|\lambda_i - \hat{\lambda}(p_i)\|T > \delta_n) \leq (c) \sum_{i=1}^{k} \sum_{j=1}^{d} \mathbb{P}(|\lambda_i^j - \hat{\lambda}^j(p_i)| > \delta_n/T) \leq (d) \frac{C_4}{n^\eta},
\]

where \( C_4 > 0 \) is suitable large, (a), (b), (c) follow from union bounds and (d) follows from a direct application of Lemma 10 and the appropriate choice of \( C_1 \).

Hence,

\[
n \sum_{i=1}^{k} p_i \cdot \lambda_i \mathbb{E}[\hat{t}_i] \geq n \left[ V_{(P_0)} - (C_2 + C_3) \max\{\delta_n, \tau_n\} \right] \left(1 - \frac{C_4}{n^\eta}\right). \tag{C.1-4}
\]

We now look into the penalty term, i.e., the second term on the RHS of (C.1-1). To that end, let \( C' > 0 \) to be a constant to be specified, \( \delta'_n = C'\delta_n \) and put
\[\mathcal{E} := \left\{ \omega : Y_n - nx \leq n\delta'_n \right\}\]

and note that

\[
\mathbb{E}\left[(Y_n - nx)^+ \right] = \mathbb{E}\left[(Y_n - nx)^+ \mid \mathcal{E}\right]\mathbb{P}(\mathcal{E}) + \mathbb{E}\left[(Y_n - nx)^+ \mid \mathcal{E}^c\right]\mathbb{P}(\mathcal{E}^c) \\
\leq n\delta'_n e + \mathbb{E}\left[(Y_n - nx)^+ \mid \mathcal{E}^c\right]\mathbb{P}(\mathcal{E}^c) \\
\leq^{(a)} n\delta'_n e + (n\delta'_n + 1 + nMT)\mathbb{P}(\mathcal{E}^c)e,
\]

where \((a)\) follows from the definition of \(\mathcal{E}\) and the fact that for a Poisson random variable \(Z\) with mean \(\mu\), \(\mathbb{E}[Z \mid Z > a] \leq a + 1 + \mu\). Now,

\[
\mathbb{P}(\mathcal{E}^c) = \mathbb{P}\left(\sum_{i=1}^{k} A\left[N\left(X_n^{(L)} + X_n^{(i)}\right) - N\left(X_n^{(L)} + X_n^{(i-1)}\right)\right] + \sum_{i=1}^{k} An\tilde{\lambda}(p_i)\Delta_n \not\leq nx + n\delta'_n\right) \\
\leq \mathbb{P}\left(\sum_{i=1}^{k} A\left[N\left(X_n^{(L)} + X_n^{(i)}\right) - N\left(X_n^{(L)} + X_n^{(i-1)}\right)\right] \not\leq nx + \frac{1}{2}n\delta'_n\right) \\
+ \mathbb{P}\left(\sum_{i=1}^{k} An\tilde{\lambda}(p_i)\Delta_n \not\leq \frac{1}{2}n\delta'_n\right).
\]

Using Lemma 10, the second term on the RHS of (C.1-5) is seen to be bounded by \(C_5/n^7\). On the other hand, the first term on the RHS of (C.1-5) can be bounded as follows

\[
\mathbb{P}\left(\sum_{i=1}^{k} A\left[N\left(X_n^{(L)} + X_n^{(i)}\right) - N\left(X_n^{(L)} + X_n^{(i-1)}\right)\right] \not\leq nx + \frac{1}{2}n\delta'_n\right) \\
\leq \mathbb{P}\left(\sum_{i=1}^{k} A\left[N\left(X_n^{(L)} + X_n^{(i)}\right) - N\left(X_n^{(L)} + X_n^{(i-1)}\right)\right] - n\tilde{\lambda}(p_i)\hat{\ell}_i \not\leq \frac{1}{4}n\delta'_n\right) \\
+ \mathbb{P}\left(\sum_{i=1}^{k} An(\lambda_i - \tilde{\lambda}(p_i))\hat{\ell}_i \not\leq \frac{1}{4}n\delta'_n\right) + \mathbb{P}\left(\sum_{i=1}^{k} An\tilde{\lambda}(p_i)\hat{\ell}_i \not\leq nx\right) \tag{C.1-6}
\]

Note that the feasibility of \(\hat{\ell}\) for \((\hat{P})\) implies that the last term on the RHS above is equal to zero. With an appropriate choice of \(C''\), Lemma 10 yields that the first
two terms on the RHS of (C.1-6) are bounded by \( C_6/n^7 \) for \( C_6 > 0 \) suitably large.

We deduce that

\[
\mathbb{E}\left[(Y_n - nx)^+\right] \leq n \delta_n' + (n \delta_n' + 1 + nM) \frac{C_5 + C_6}{n^7} e,
\]

Combining the above with (C.1-1) and (C.1-4), we have

\[
J_n^\pi \geq n \sum_{i=1}^k p_i \cdot \lambda_i \mathbb{E}\left[\tilde{h}_i\right] - \tilde{M} e \cdot \mathbb{E}\left[(Y_n - nx)^+\right] \\
\geq n \left[V_{(F_0)}^* - (C_2 + C_3) \max\{\delta_n, \tau_n\}\right] \left(1 - \frac{C_4}{n^7}\right) \\
- \tilde{M} [n \delta_n' + (n \delta_n' + 1 + nM)] \frac{C_5 + C_6}{n^7} \\
\geq n V_{(F_0)}^* - C_9 n \max\{\delta_n, \tau_n\} + 1/n^7).
\]

**Step 3.** We now conclude the proof. Recalling that \( m^D > 0 \) bounds below \( V_{(F_0)}^* \) for all \( \lambda \in \mathcal{L} \), we have

\[
\frac{J_n^\pi}{J_n^*} \geq \frac{J_n^\pi}{n V_{(F_0)}^*} \geq 1 - \frac{C_9 (\max\{\delta_n, \tau_n\} + 1/n^7)}{m^D}
\]

implying that uniformly over \( \lambda \in \mathcal{L} \)

\[
\liminf_{n \to \infty} \frac{J_n^\pi}{J_n^*} \geq 1.
\]

This, in conjunction with the inequality \( J_n^\pi \leq J_n^* \), completes the proof.

To obtain the rate of convergence stated in (4.3.6) in Remark 2 note that the orders of the terms \( \delta_n \) and \( \tau_n \) are balanced by choosing \( \tau_n \propto n^{-1/3} \). With this choice we have

\[
\sup_{\lambda \in \mathcal{L}} \limsup_{n \to \infty} \frac{1 - J_n^\pi/J_n^*}{(\log n)^{1/2} n^{-1/3}} < \infty.
\]
Proof of Theorem 5. Fix $\lambda \in \mathcal{L}$ and $\eta \geq 2$. For simplicity, we restrict attention to the product set $\mathcal{D}_p = \prod_{i=1}^d [p_i^*, p_i^*]$. Let $\bar{M} = \max_{1 \leq i \leq d} p^i$ be the maximum price a customer will ever pay for a product. It is easy to verify that the deterministic optimization problem given (4.4.2) is a convex problem whose solution is given by a constant price vector $\bar{p}$ (cf. Gallego and van Ryzin (1997)). Let $\pi$ be the policy defined by means of Algorithm 6.

Step 1. We first focus on the the learning and optimization phases. Let $\tau_n$ be such that $\tau_n = o(1)$ and $n \tau_n \to \infty$ at a polynomial rate. Let $\kappa_n$ be a sequence of integers such that $\kappa_n \to \infty$ and $n \Delta_n := n \tau_n / \kappa_n \to \infty$ at a polynomial rate. Divide each interval $[p_i^*, p_i^*]$, $i = 1, \ldots, d$ into $\lceil \kappa_n^{1/d} \rceil$ equal length intervals and consider the resulting grid in $\mathcal{D}_p$. The latter has $\kappa'_n = \lceil \kappa_n^{1/d} \rceil^d$ hyper rectangles. For each one, let $p_i$ be the largest vector (where the largest vector of a hyper rectangle $\prod_{i=1}^d [a_i, b_i]$ is defined to be $(b_1, \ldots, b_d)$) and consider the set $P^{\kappa'_n} = \{p_1, p_2, \ldots, p_{\kappa'_n}\}$. Note that $\kappa'_n / \kappa_n \to 1$ as $n \to \infty$ and with some abuse of notation, we use both $\kappa_n$ and $\kappa'_n$ interchangeably.

Now partition $[0, \tau_n]$ into $\kappa_n$ intervals of length $\Delta_n$ and apply the price vector $p_i$ on the $i^{th}$ interval. Define

$$\hat{\lambda}(p_i) = \frac{N \left( n \Delta_n \sum_{j=1}^{i-1} \lambda(p_j) \right) - n \Delta_n \sum_{j=1}^{i-1} \lambda(p_j)}{n \Delta_n}, \quad i = 1, \ldots, \kappa_n,$$

where $N(\cdot)$ is the $d$-vector of unit rate Poisson processes. Thus $\hat{\lambda}(p_i)$ denotes the number of requests for each product over successive intervals of length $\Delta_n$,
We now choose the “best” price among “almost feasible prices.” Specifically, we let \( \delta_n = C_1 (\log n)^{1/2} \max\{1/\kappa_n, (n \Delta_n)^{-1/2}\} \) with \( C_1 = 2 \max\{1, \bar{p}\} C(\eta) \) where \( C(\eta) \) is defined in Lemma 10. Set \( \hat{r}(p_i) = p_i \cdot \hat{\lambda}(p_i) \) if \( A\hat{\lambda}(p_i) I \leq x + e \delta_n \); otherwise set \( \hat{r}(p_i) = 0 \). The objective of this step is to discard solutions of the deterministic problem which are essentially infeasible. (The slack term \( \delta_n \) allows for “noise” in the observations.) Let
\[
\hat{p} = p_i^* \quad \text{where} \quad i^* = \arg \max \{\hat{r}(p_i), i = 1, \ldots, \kappa_n\}. \tag{C.1-7}
\]

**Step 2.** Here, we derive a lower bound on the expected revenues under the policy \( \pi \). We will need the following lemma whose proof is deferred to Appendix C.2.

**Lemma 12** Let \( P_f^m = \{p_i \in P^{\kappa_n} : A\hat{\lambda}(p_i) I \leq x + \delta_n e\} \). Then for a suitably large constant \( C_3 > 0 \)
\[
\Pr\left( r(\tilde{p}) - r(\hat{p}) > \delta_n \right) \leq \frac{C_3}{n^\eta},
\]
\[
\Pr\left( \hat{p} \notin P_f^m \right) \leq \frac{C_3}{n^\eta}.
\]

We define \( X_n^{(L)} = \sum_{i=1}^{\kappa_n} \lambda(p_i)n \Delta_n, X_n^{(P)} = \lambda(\hat{p})n(T - \tau_n) \) and put \( Y_n = AN(X_n^{(L)} + X_n^{(P)}) \). In the rest of the proof, we will use the fact that given \( \hat{p}, Y_n = \sum_{i=1}^{\kappa_n} A\hat{\lambda}(p_i)n \Delta_n + AN(X_n^{(L)} + X_n^{(P)}) - AN(X_n^{(L)}) \) and that \( N(X_n^{(L)} + X_n^{(P)}) - N(X_n^{(L)}) \) has the same distribution as \( N(X_n^{(P)}) \). Recalling Comment 1 in the preamble of the appendix, note that \( Y_n \) is the total potential demand (for
resources) under \( \tau \) if one would never use \( p_\infty \) and that one can lower bound the revenues under \( \tau \) as follows

\[
J_n^\tau \geq \mathbb{E}\left[ \hat{p} \cdot \left( N(X_n^{(L)} + X_n^{(P)}) - N(X_n^{(L)}) \right) \right] - \bar{M}e \cdot \mathbb{E}\left[ \left( Y_n - nx \right)^+ \right]. \tag{C.1-8}
\]

The first term on the RHS of (C.1-8) can be bounded as follows

\[
\mathbb{E}\left[ \hat{p} \cdot \left( N(X_n^{(L)} + X_n^{(P)}) - N(X_n^{(L)}) \right) \right] = \mathbb{E}\left[ \mathbb{E}\left[ \hat{p} \cdot N(\lambda(\hat{p})n(T - \tau_n)) \mid \hat{p} \right] \right] = \mathbb{E}\left[ r(\hat{p}) \right] n(T - \tau_n)
\]

\( \overset{(a)}{=} \left\{ r(\hat{p}) + \mathbb{E}\left[ r(\hat{p}) - r(\hat{p}) \mid r(\hat{p}) - r(\hat{p}) > -\delta_n \right] \mathbb{P}\left( r(\hat{p}) - r(\hat{p}) > -\delta_n \right) \right. \]

\[ + \mathbb{E}\left[ r(\hat{p}) - r(\hat{p}) \mid r(\hat{p}) - r(\hat{p}) \leq -\delta_n \right] \mathbb{P}\left( r(\hat{p}) - r(\hat{p}) \leq -\delta_n \right) \left\} n(T - \tau_n) \]

\( \overset{(b)}{=} \left[ r(\hat{p}) - \delta_n - \frac{C_4}{n^{\eta}} \right] n(T - \tau_n), \tag{C.1-9} \]

where \( C_4 \) is a suitably large positive constant. Note that (a) follows from conditioning and (b) follows from Lemma 12 and the fact that \( r(\cdot) \) is bounded say by \( d\bar{M}M \). Let us now examine the second term on the RHS of (C.1-8). Let \( C' > 0 \) be a constant to be specified later and \( \delta'_n = C' \delta_n \).

\[
\mathbb{E}\left[ \left( Y_n - nx \right)^+ \right] = \mathbb{E}\left[ \left( Y_n - nx \right)^+ \mid Y_n - nx \leq n\delta'_n e \right] \mathbb{P}(Y_n - nx \leq n\delta'_n e)
\]

\[ + \mathbb{E}\left[ \left( Y_n - nx \right)^+ \mid Y_n - nx \not\leq n\delta'_n e \right] \mathbb{P}(Y_n - nx \not\leq n\delta'_n e) \leq n\delta'_n e + \mathbb{E}\left[ Y_n \mid Y_n \not\leq nx + n\delta'_n e \right] \mathbb{P}(Y_n - nx \not\leq n\delta'_n e), \]

Now, for a Poisson random variable \( Z \) with mean \( \mu \), it is easy to see that \( \mathbb{E}[Z \mid Z > a] \leq a + 1 + \mu \). In particular, each component of \( Y_n \) is a Poisson random variable
with rate less than $nMT$ and hence

$$
\mathbb{E}[Y_n | Y_n \leq nx + n\delta'_n e] \leq nx + (n\delta'_n + 1 + nMT)e.
$$

Let us evaluate the probability to run out of some resource by more than $n\delta'_n$.

Specifically,

$$
P(Y_n - nx \leq n\delta'_n e)
\leq P\left(AN(n\lambda(\hat{p})(T - \tau_n)) - A_n\lambda(\hat{p})(T - \tau_n) \leq \frac{1}{3}n\delta'_n e\right)
\leq \mathbb{E}\left[P\left(AN(n\lambda(\hat{p})(T - \tau_n)) - A_n\lambda(\hat{p})(T - \tau_n) \leq \frac{1}{3}n\delta'_n e\right)\right]
\leq \frac{C_3}{n^n}.
$$

Consider the first term on the RHS of (C.1-10). We have

\begin{align*}
n\delta'_n &> n(T - \tau_n)3C(\eta)(\log n)^{1/2}(n(T - \tau_n))^{-1/2} \text{ for } n \text{ large enough and hence, if } C' \geq 3T, \text{ one can condition on } \hat{p} \text{ and apply Lemma 10 (with } \mu = \lambda(\hat{p}), \tau_n = n(T - \tau_n) \text{ to get}

P\left(AN(\lambda(\hat{p})n(T - \tau_n)) - A\lambda(\hat{p})n(T - \tau_n) \leq \frac{1}{3}n\delta'_n e\right)
\leq \mathbb{E}\left[P\left(AN(\lambda(\hat{p})n(T - \tau_n)) - A\lambda(\hat{p})n(T - \tau_n) \leq \frac{1}{3}n\delta'_n e\right)\right]
\leq \frac{C_3}{n^n}.
\end{align*}

Consider now the second term on the RHS of (C.1-10)

\begin{align*}
P\left(A\lambda(\hat{p})n(T - \tau_n) \leq n(x + \frac{\delta'_n}{3} e)\right)
= P\left(A[\lambda(\hat{p})T - \hat{\lambda}(\hat{p})] + A\hat{\lambda}(\hat{p})T \leq \frac{1}{1 - \tau_n/T}(x + \frac{\delta'_n}{3} e)\right)
\leq P\left(A[\lambda(\hat{p})T - \hat{\lambda}(\hat{p})] \leq \frac{\delta'_n}{6} e\right) + P\left(A\hat{\lambda}(\hat{p})T \leq x + \frac{\delta'_n}{6} e\right)
= P\left(A(\lambda(\hat{p})n\Delta_n T - \hat{\lambda}(\hat{p})n\Delta_n T) \leq n\Delta_n \frac{\delta'_n}{6} e\right) + P\left(A\hat{\lambda}(\hat{p})T \leq x + \frac{\delta'_n}{6} e\right). \quad (C.1-11)
\end{align*}
Suppose that $C' \geq 6$. Then by Lemma 12, the second term above is bounded by $C_5/n^\eta$ for a large enough choice of $C_5 > 0$. The first term on the RHS of (C.1-11) is upper bounded by $C_3/n^\eta$ by Lemma 10. Consider the third term on the RHS of (C.1-10).

\[
\mathbb{P}\left( \sum_{i=1}^{\kappa_n} A\lambda(p_i)n\Delta_n \not\leq \frac{1}{3} n\delta'_n e \right)
\leq \sum_{i=1}^{\kappa_n} \mathbb{P}\left( A\lambda(p_i)n\Delta_n \not\leq \frac{1}{3} n\delta'_n e \right)
= \sum_{i=1}^{\kappa_n} \mathbb{P}\left( A[N(\lambda(p_i)n\Delta_n) - \lambda(p_i)n\Delta_n] \not\leq n\Delta_n \left( \frac{1}{3} \delta'_n e - A\lambda(p_i) \right) \right).
\]

Now if $\delta'_n/\tau_n \to \infty$ (which holds, for example if $\tau_n = n^{-1/(d+3)}$, $\kappa_n = n^{d/(d+3)}$), then for $n$ sufficiently large, we have $(1/3)\delta'_n/\tau_n e - A\lambda(p_i) \geq 1$ for all $i = 1, \ldots, \kappa_n$ and Lemma 10 yields

\[
\mathbb{P}\left( \sum_{i=1}^{\kappa_n} A\lambda(p_i)n\Delta_n \not\leq \frac{1}{3} n\delta'_n e \right) \leq \frac{\kappa_n C_3 n^\eta}{n^{\eta-1}} \leq \frac{C_3}{n^{\eta-1}}.
\]

We conclude that with $C'' = \max\{3T, 6\}$ and for some $C_6 > 0$, $\mathbb{P}(Y_n \not\leq nx + n\delta'_n e) \leq C_6/n^{\eta-1}$, and in turn

\[
\mathbb{E}\left[ (Y_n - nx)^+ \right] \leq n\delta'_n e + \mathbb{E}\left[ Y_n \mid Y_n \not\leq nx + n\delta'_n e \right] \frac{C_6}{n^{\eta-1}}. \quad \text{(C.1-12)}
\]

Combining (C.1-8), (C.1-9) and (C.1-12) we have

\[
J_n^\pi \geq \left[ r(\tilde{p}) - \delta_n - \frac{C_4}{n^\eta} \right] n(T - \tau_n) - \tilde{M}n\delta'_n - \tilde{M}(nx \cdot e + n\delta'_n + 1 + nMT) \frac{C_6}{n^{\eta-1}}
= r(\tilde{p})nT - n\left[ (T - \tau_n)\delta_n + (T - \tau_n) \frac{C_4}{n^\eta} + \tilde{M}C'\delta_n \right.
+ (\tilde{M}x \cdot e + MT) \frac{C_6}{n^{\eta-1}} + C'\delta_n \frac{C_6}{n^{\eta-1}} + \frac{C_6}{n^{\eta-2}}] \\
\quad \overset{(a)}{\geq} r(\tilde{p})nT - nC_7 \left[ \tau_n + \delta_n + 1/n^{\eta-2} \right],
\]
where (a) follows from the fact that $\delta_n \to 0$ and by choosing $C_7 > 0$ is suitably large.

**Step 3.** We now conclude the proof. Note that under the current assumptions, $\mathcal{D}_\lambda$ is convex. Gallego and van Ryzin (1997, Theorem 1) show that under these conditions the optimal value of problem (4.4.2) say $J_n^D$ serves as upper bound to $J_n^*$. Note that $J_n^D = nr(p)T$. Define $f(n) := C_7 \left[ \tau_n + \delta_n + 1/n^{\eta-2} \right]$ and note that $f(n) \geq 0$ for all $n \geq 0$ and that $f(n) \to 0$ as $n \to \infty$. In addition $f(n)$ does not depend on the specific underlying demand $\lambda$. By the remark in the preamble, $J_n^D \geq nm^D > 0$ and hence

$$\frac{J_n^\pi}{J_n^*} \geq \frac{J_n^\pi}{J_n^D} \geq 1 - \frac{f(n)}{m^D}$$

implying that uniformly over $\lambda \in \mathcal{L}$

$$\liminf_{n \to \infty} \frac{J_n^\pi}{J_n^*} \geq 1.$$

This, in conjunction with the inequality $J_n^\pi \leq J_n^*$, completes the proof.

To obtain the rate of convergence stated in (4.4.4) in Remark 1 note that the orders of the terms $\tau_n$ and $\delta_n$ are balanced by choosing $\tau_n = n^{-1/(d+3)}$ and $\kappa_n = n^{d/(d+3)}$. With this choice we have for $C_8 = C_7/m^D$,

$$f(n)/m^D = C_8[(\log n)^{1/2}/n^{1/(d+3)} + 1/n^{\eta-1}],$$

implying that

$$\sup_{\lambda \in \mathcal{L}} \limsup_{n \to \infty} \frac{1 - J_n^\pi / J_n^*}{(\log n)^{1/2} n^{-1/(d+3)}} < \infty.$$ 

**Proof of Theorem 6.** The proof follows three steps. In the first step, we bound the error between a function $\lambda \in \mathcal{L}$ and an approximation based on the observa-
tions available. The latter is done using local polynomials along the lines provided in Nemirovski (2000, Chap. 1). The main difference is that the noise associated with observations of the demand function is the deviation of Poisson increments from their mean rather than Gaussian random variables. In the second step, we bound below the expected revenues achieved by the proposed policy and the last step concludes with balancing all error sources.

**Step 1.** Choose the sequences \( \tau_n, \kappa_n \) and \( P^{\kappa_n} = \{p_1, \ldots, p_{\kappa_n}\} \) as in Step 1 in the proof of Theorem 5. Let

\[
z_i = \frac{N\left( n\Delta_n \sum_{j=1}^{i} \lambda(p_j) \right) - N\left( n\Delta_n \sum_{j=1}^{i-1} \lambda(p_j) \right)}{n\Delta_n}, \quad i = 1, \ldots, \kappa_n,
\]

where \( N(\cdot) \) is the \( d \)-vector of unit rate Poisson processes. Thus \( z_i \) denotes the number of requests for each product over successive intervals of length \( \Delta_n \), normalized by \( n\Delta_n \).

Let us focus on the first component of the demand function which we denote by \( f(\cdot) \) to simplify notation \((f(\cdot) := \lambda^1(\cdot))\). Let \( y \) denote the vector \((z_1^1, \ldots, z_{\kappa_n}^1)\). Let \( h_n = o(1) \) such that \( h_n\kappa_n^{1/d} \geq s + 1 \).

We provide below some properties that the weights defined in (4.5.2) satisfy. In Nemirovski (2000, Lemma 1.3.1), it is established that

\[
\gamma(p) = \sum_{i: p_i \in B_p} \omega_i^B(p) \gamma(p_i) \quad \text{for every polynomial } \gamma \text{ of degree } k(C.1-13)
\]

\[
\| \omega^B(p) \|_2^2 := \sum_{i: p_i \in B_p} (\omega_i^B(p))^2 \leq \frac{C_1}{\kappa_n h_n^d}, \quad (C.1-14)
\]

\[
\| \omega^B(p) \|_1 := \sum_{i: p_i \in B_p} |\omega_i^B(p)| \leq C_1, \quad (C.1-15)
\]
for some positive constant $C_1 > 0$. In other words, one is able to reproduce the
value of any polynomial of degree $k$ through its value at the points in $G$ and the
weights $\omega^B(p)$. In addition, one is able to control uniformly the norms of the
weights. We now define an approximation for the function $f(\cdot)$ as follows

$$\hat{f}(p; y) = \sum_{i: p_i \in B_p} \omega_i^B(p)y_i.$$  \hfill (C.1-16)

In what follows we bound the difference between the function $f(\cdot)$ and its approx-
imation $\hat{f}(p; y)$. Let $\theta(p)$ be a Taylor expansion of order $k$ of $f(p)$ around a point
in $B_p$.

$$\left| f(p) - \hat{f}(p; y) \right|$$

$$= \left| f(p) - \sum_{i: p_i \in B_p} \omega_i^B(p)y_i \right|$$

$$= \left| f(p) - \sum_{i: p_i \in B_p} \omega_i^B(p)[\theta(p_i) + f(p_i) - \theta(p_i) + y_i - f(p_i)] \right|$$

$$\leq \left| f(p) - \theta(p) \right| + \left| \sum_{i: p_i \in B_p} \omega_i^B(p)(f(p_i) - \theta(p_i)) \right| + \left| \sum_{i: p_i \in B_p} \omega_i^B(p)(y_i - f(p_i)) \right|$$

$$\leq \sup_{q \in B_p} \left| f(q) - \theta(q) \right| \left[ 1 + \sum_{i: p_i \in B_p} |\omega_i^B(p)| \right] + \left| \sum_{i: p_i \in B_p} \omega_i^B(p)(y_i - f(p_i)) \right|, \hfill (C.1-17)$$

where (a) follows from the fact that $\theta(\cdot)$ is a polynomial of degree $k$, the property
(C.1-13) and the triangular inequality.

Let $\xi_i = y_i - f(p_i)$ and $\zeta^B_p = \frac{1}{\|w_p^B\|_2} \sum_{i: p_i \in B_p} \omega_i^B(p)\xi_i$ and $\Theta_n = \sup_{p \in \mathcal{D}_n} |\zeta^B_p|$. Note that (4.5.2) implies that every component $\omega_i^B(p)$ is a polynomial in $p$ of degree
less or equal than $k$ and hence can be written as $\omega_i^B(p) = \sum_{j: p_j \in B_p} \omega_j^B(p)\omega_i^B(p_j)$. 

Now, we have

\[
\sup_{p \in \mathcal{D}_p} |\zeta^B_p| = \sup_{p \in \mathcal{D}_p} \frac{1}{\|\omega^B(p)\|_2} \left| \sum_{i:p_i \in B_p} \omega^B_i(p) \xi_i \right|
\]

\[
= \sup_{p \in \mathcal{D}_p} \frac{1}{\|\omega^B(p)\|_2} \left| \sum_{i:p_i \in B_p} \sum_{j:p_j \in B_p} \omega^B_j(p) \omega^B_i(p_j) \xi_i \right|
\]

\[
\leq (a) \sup_{p \in \mathcal{D}_p} \frac{1}{\|\omega^B(p)\|_2} \left[ \sum_{j:p_j \in B_p} \left( \sum_{i:p_i \in B_p} \omega^B_i(p_j) \xi_i \right)^2 \right]^{1/2}
\]

\[
= \sup_{p \in \mathcal{D}_p} \left[ \sum_{j:p_j \in B_p} \left( \sum_{i:p_i \in B_p} \omega^B_i(p_j) \xi_i \right)^2 \right]^{1/2},
\]

where \(a\) follows from Cauchy-Schwarz inequality. Let

\[
\beta_j = \frac{1}{\|\omega^B(p_j)\|_2} \sum_{i:p_i \in B_p} \omega^B_i(p_j) \xi_i.
\]

Note that as \(p\) covers \(\mathcal{D}_p\), there are only \(n' \leq \kappa_n^2\) possible values for

\[
\left[ \sum_{j:p_j \in B_p} \left( \sum_{i:p_i \in B_p} \omega^B_i(p_j) \xi_i \right)^2 \right]^{1/2}.\]

Let \(k = 1, ..., n'\) index those random variables and denote by \(\gamma_k\) the \(k^{th}\) possible value. Note that for some \(p\),

\[
\gamma_k := \left[ \sum_{j:p_j \in B_p} \left( \sum_{i:p_i \in B_p} \omega^B_i(p_j) \xi_i \right)^2 \right]^{1/2} = \left[ \sum_{j:p_j \in B_p} \|\omega^B(p_j)\|_2^2 \beta_j^2 \right]^{1/2}.
\]

Let \(u_n = (C \log n)^{1/2} (n \Delta_n)^{-1/2}\) where \(C\) is a constant to be defined. Let \(\alpha > 0\)
and define $\alpha'_j = \alpha / \|\omega^B(p_j)\|_2$. We have

$$\mathbb{P}(\beta_j > u_n)$$

\[ \leq \exp(-\alpha u_n) \mathbb{E}[\exp(\alpha_0)] \]

\[ = \exp(-\alpha u_n) \mathbb{E}\left[ \exp\left( \alpha'_j \sum_{i:p_i \in B_p} \omega_i^B(p_j) \xi_i \right) \right] \]

\[ = \exp(-\alpha u_n) \prod_{i:p_i \in B_p} \mathbb{E}[\exp(\alpha'_j \omega_i^B(p_j) \xi_i)] \]

\[ = \exp(-\alpha u_n) \prod_{i:p_i \in B_p} \exp\left(-\alpha'_j \omega_i^B(p_j) f(p_i) \exp\left( f(p_i) n \Delta_n \left[ \exp\left((\alpha'_j / n \Delta_n) \omega_i^B(p_j)\right) - 1 \right]\right) \]

\[ \leq \exp(-\alpha u_n) \prod_{i:p_i \in B_p} \exp\left(\Lambda n \Delta_n (3/2) \left( \left(\alpha'_j / n \Delta_n\right) \omega_i^B(p_j) \right)^2 \right) \]

\[ = \exp(-\alpha u_n) \exp\left(\Lambda (3/2) \alpha^2 (n \Delta_n)^{-1}\right) \]

where (a) follows from the Chernoff bound, (b) follows from the fact that $\exp(x) - 1 \leq x + (3/2)x^2$ as long as $x \leq 1$, that $\alpha$ will be chosen to shrink to zero and that $\lambda(p_i) \leq \Lambda$. Now choosing $\alpha = (1/(3\Lambda))u_n n \Delta_n$, we obtain

$$\mathbb{P}(\beta_j > u_n) \leq \exp(-6\Lambda)^{-1} u_n^2 n \Delta_n$$

$$\leq \exp(-6\Lambda)^{-1} C \log n).$$

Similarly,

$$\mathbb{P}(\beta_j < -u_n) \leq \exp(-6\Lambda)^{-1} C \log n).$$
Now,

\[
P\left( \gamma_k > C_1^{1/2}u_n \right) \leq P\left( \sum_{j:p_j \in B_p} \left\| \omega^B(p_j) \right\|_2^2 \beta_j^2 > C_1 u_n^2 \right)
\]

\[
\leq P\left( \sum_{j:p_j \in B_p} \kappa_n^{-1} h_n^{-d} \beta_j^2 > u_n^2 \right)
\]

\[
\leq \sum_{j:p_j \in B_p} \left[ P(\beta_j > u_n) + P(\beta_j < -u_n) \right]
\]

\[
\leq 2\kappa_n h_n^d \exp(-(6\lambda)^{-1}C \log n),
\]

where (a) follows from (C.1-14). Focusing on \( \Theta_n \), we have

\[
P\left( \Theta_n > C_1^{1/2}u_n \right) \leq n'P\left( \gamma_k > C_1^{1/2}u_n \right)
\]

\[
\leq n'2\kappa_n h_n^d \exp(-(6\lambda)^{-1}C \log n)
\]

\[
\leq 2\kappa_n h_n^d \exp(-(6\lambda)^{-1}C \log n).
\]

By choosing \( C \) sufficiently large, we have

\[
P\left( \Theta_n > u_n \right) \leq \frac{1}{n^2}.
\]

Coming back to (C.1-17), and noting that by the assumptions on \( f \), the difference between \( \theta(p) \) and \( f(p) \) is uniformly bounded by \( C_2 h_n^s \) on \( B_p \), we have

\[
\sup_{p \in D_p} \left| f(p) - \widehat{f}(p; y) \right| \leq (a) C_2 h_n^s + \kappa_n^{-1/2} (h_n)^{-d/2} u_n \frac{\Theta_n}{u_n}
\]

\[
\leq (b) C_2 h_n^s + (n\tau_n)^{-1/2} (h_n)^{-d/2} (C \log n)^{1/2} \frac{\Theta_n}{u_n},
\]

where (a) follows from (C.1-14) and (b) follows from the definition of \( u_n \) and \( \Delta_n = \tau_n / \kappa_n \). The choice \( h_n = (n\tau_n)^{-1/(2s+d)} \) balances the error terms above and with such a choice, we have for some \( C_3 > 0 \),

\[
\sup_{p \in D_p} \left| f(p) - \widehat{f}(p; y) \right| \leq C_3 (\log n)^{1/2} (n\tau_n)^{-s/(2s+d)} \left[ 1 + \frac{\Theta_n}{u_n} \right].
\]
We have just established that

**Lemma 13** Suppose $\kappa_n \geq (s + 1)(n\tau_n)^{1/(2s+d)}$, then following Step 1 of Algorithm 6, one can construct an estimate of the demand function $\tilde{\lambda}(\cdot; y)$ such that for some $C_4 > 0$, for all $n \geq 1$,

$$
P \left( \sup_{\lambda \in \mathcal{L}} \sup_{p \in \mathcal{D}_p} \|\tilde{\lambda}(p; y) - \lambda(p)\|_\infty > C_4 \frac{(\log n)^{1/2}}{(n\tau_n)^{s/(2s+d)}} \right) \leq \frac{C_4}{n^2}. \quad (C.1-18)
$$

**Steps 2 and 3.** Following a similar reasoning as the one in Steps 2 and 3 in the proof of Theorem 5, one arrives at the following inequality

$$
\frac{J^*_n}{J^*_n} \geq 1 - C_5 [\tau_n + \delta_n + 1/n^{\eta-1}],
$$

where $C_5$ is a positive constant and $\delta_n = (\log n)^{1/2}(n\tau_n)^{-s/(2s+d)}$ and $\eta \geq 2$ is fixed.

The choice of $\tau_n \approx n^{-s/(3s+d)}$ leads to

$$
\sup_{\lambda \in \mathcal{L}} \limsup_{n \to \infty} \frac{1 - J^*_n/J^*_n}{(\log n)^{1/2}n^{-1/(3+d/s)}} < \infty.
$$

**C.2 Proofs of Auxiliary Results**

In what follows $C_1', i \geq 1$ will denote positive constants that depend only on $A, x, T$ and the parameters of the class $\mathcal{C}$, but not on a specific function $\lambda \in \mathcal{L}$.

**Proof of Lemma 10.** Let $\mathcal{J}_i = \{j \in \{1, \ldots, d\} : a_{ij} \neq 0\}$. We proceed with the
following inequalities

\[
\mathbb{P}\left(A[N(\mu r_n) - \mu r_n] \leq r_n \epsilon_n \right) \quad \text{(C.2-1)}
\]

\[
\leq \sum_{i=1}^{\ell} \mathbb{P}\left(\sum_{j=1}^{d} a_{ij} [N(\mu_j r_n) - \mu_j r_n] > r_n \epsilon_n \right)
\]

\[
\leq \sum_{i=1}^{\ell} \sum_{j \in J_i} \mathbb{P}\left( N(\mu_j r_n) - \mu_j r_n > \frac{r_n \epsilon_n}{d a_{ij}} \right)
\]

\[
\leq \ell \sum_{j=1}^{d} \mathbb{P}\left( N(\mu_j r_n) - \mu_j r_n > \frac{r_n \epsilon_n}{d a} \right)
\]

\[
\leq \ell \sum_{j=1}^{d} \exp\left\{-\theta_j r_n (\mu_j + \frac{\epsilon_n}{d a}) + \left(\exp\{\theta_j\} - 1\right) \mu_j r_n \right\}, \quad \text{(C.2-2)}
\]

where (a) follows from a union bound and (b) follows from the Chernoff bound.

The expression in each of the exponents is minimized for the choice of \( \theta_j > 0 \) defined by

\[
\theta_j = \log\left(1 + \frac{\epsilon_n}{d a \mu_j}\right). \quad \text{(C.2-3)}
\]

Plugging back into (C.2-2) yields

\[
\mathbb{P}\left(A[N(\mu r_n) - \mu r_n] \leq r_n \epsilon_n \right)
\]

\[
\leq \ell \sum_{j=1}^{d} \exp\left\{-\log\left(1 + \frac{\epsilon_n}{d a \mu_j}\right) (\mu_j + \frac{\epsilon_n}{d a \mu_j}) + \frac{\epsilon_n}{d a} \right\}
\]

\[
\leq \ell d \exp\left\{r_n \left(-\log\left(1 + \frac{\epsilon_n}{d a M}\right) (M + \frac{\epsilon_n}{d a}) + \frac{\epsilon_n}{d a} \right) \right\}. \quad \text{(C.2-4)}
\]

For the last inequality, note that the derivative of the term in the exponent with respect to \( \mu_j \) is given by \(-\log(1 + \epsilon_n/\mu_j) + \epsilon_n/\mu_j\), which is always positive for \( \epsilon_n > 0 \). Now, using a Taylor expansion we get that for some \( \xi \in [0, \frac{\epsilon_n}{d a M}] \),

\[
-M \left[ \log\left(1 + \frac{\epsilon_n}{d a M}\right) (1 + \frac{\epsilon_n}{d a M}) - \frac{\epsilon_n}{d a M} \right] = -\frac{1}{2} \frac{\epsilon_n}{1 + \xi d^2 a^2 M} \leq -\frac{\epsilon_n^2}{4d^2 a^2 M},
\]
where the last inequality holds only if $\varepsilon_n/(d\bar{a}M) \leq 1$ (which is valid for sufficiently large $n$). Substituting for $\varepsilon_n$, we get

$$
p\left( A[N(\mu r_n) - \mu r_n] \leq r_n \varepsilon_n \right) \leq \ell d \exp\left\{ -\frac{(C(\eta))^2 \log n}{4d^2 \bar{a}^2 M} \right\}
$$

$$= \frac{\ell d}{n^\gamma},$$

Hence the first result follows. The other inequality goes through in a similar fashion. This completes the proof.

**Proof of Lemma 11.** Suppose first that for all $i = 1, \ldots, k$, $A\lambda_i \hat{t}_i \leq x/k$, then

$$\sum_{i=1}^{k} A\lambda_i \hat{t}_i \leq x,$$

Suppose now that there exists $i^*, 1 \leq i^* \leq k$ such that $A\lambda_i \hat{t}_{i^*} > x/k$. Note that this implies that $\hat{t}_{i^*} > x/(kM\|Ae\|)$ where $M$ is the constant that bounds the demand rate at any price (see ...). Let $\hat{t}'$ be defined as follows: $\hat{t}_i' = \hat{t}_i$ for all $i \neq i^*$ and $\hat{t}_{i^*}' = (\hat{t}_{i^*} - C_1' \delta_n)^+$ with $C_1' = kT^2 \max_{1 \leq i \leq k}\{(Ae)_i/x_i\}$ and $(Ae)_i$ denotes the $i$th component of the vector $Ae$. $\hat{t}'$ is clearly feasible for $(\hat{P})$ for $n$ sufficiently large.

For $\omega \in \mathcal{H}$ we have

$$\sum_{i=1}^{k} A\lambda_i \hat{t}_i' = \sum_{i=1}^{k} A\lambda_i \hat{t}_i + \sum_{i=1}^{k} A(\lambda_i - \hat{\lambda}(p_i)) \hat{t}_i - A\lambda_{i^*} C_1' \delta_n
$$

$$\leq [a] x + \max_{1 \leq i \leq k} \|\lambda_i - \hat{\lambda}(p_i)\|T A e - \frac{x}{kT} C_1' \delta_n
$$

$$\leq [b] x + \delta_n (A e - \frac{x}{kT} C_1')
$$

$$\leq [c] x,$$

where $(a)$ follows from the feasibility of $\hat{t}$ for $(\hat{P})$ and the fact that $A\lambda_i \hat{t}_{i^*} > x/k$ implies that $A\lambda_{i^*} > x/kT$; $(b)$ follows from the fact that $\omega \in \mathcal{H}$; and $(c)$ follows
from the choice of \( C' \). We deduce that for \( \omega \in \mathcal{H} \), \( \tilde{t}' \) is feasible for \((P_0)\). In addition the revenues achieved by \( \tilde{t}' \) in \((P_0)\) can be lower bounded as follows (where \( C'_2 > 0 \) is suitable large)

\[
V_{(P_0)}(\tilde{t}') = \sum_{i=1}^{k} p_i \cdot \lambda_i \hat{t}_i \\
= \sum_{i=1}^{k} p_i \cdot \hat{\lambda}(p_i) \hat{t}_i + \sum_{i=1}^{k} p_i \cdot (\lambda_i \cdot \hat{\lambda}(p_i)) \hat{t}_i - p_i \cdot \lambda_i \cdot C'_i \delta_n \\
\geq V_{(P_j)}(\hat{t}) - dM \max_{1 \leq i \leq k} \| \lambda_i - \hat{\lambda}(p_i) \| T - MMC'_i \delta_n \\
\geq V_{(P_j)}(\hat{t}) - C'_2 \delta_n.
\]

On the other hand, consider an optimal solution \( t^* \) to \((P_0)\). We can proceed with a similar reasoning, for all \( i = 1, \ldots, k \), \( A\hat{\lambda}(p_i) t^*_i \leq x/k \), then \( \sum_{i=1}^{k} A\lambda_i \tilde{t}_i \leq x \).

Now, consider the case where for some \( i' \), \( A\hat{\lambda}(p_{i'}) t^*_i > x/k \). By the definition of \( \mathcal{H} \), we have that \( t^*_i > x/(k(M + \delta_n)\|Ae\|) \). In addition \( A\hat{\lambda}(p_{i'}) > x/(kT) \). Let \( \eta_n = \max\{\tau_n, C'_3 \delta_n\} \) with \( C'_3 = kT^2 \max_{1 \leq i \leq k} \{ (Ae)_i / x_i \} \) and define \( \tilde{t}' = t^*_i - \eta_n \) and \( \tilde{t}_i = t^*_i \) for all \( i \neq i' \). Note that for \( n \) sufficiently large \( \tilde{t}_i \geq 0 \) for \( i = 1, \ldots, k \) and \( \sum_{i=1}^{k} \tilde{t}_i \leq \sum_{i=1}^{k} t^*_i - \tau_n \leq T - \tau_n \). In addition, we have for \( \omega \in \mathcal{H} \)

\[
\sum_{i=1}^{k} A\hat{\lambda}(p_i) \tilde{t}_i = \sum_{i=1}^{k} A\hat{\lambda}(p_i) t^*_i - A\hat{\lambda}(p_{i'}) \eta_n \\
= \sum_{i=1}^{k} A\lambda_i t^*_i + \sum_{i=1}^{k} A(\hat{\lambda}(p_i) - \lambda_i) t^*_i - A\hat{\lambda}(p_{i'}) \eta_n \\
\overset{(a)}{\leq} x + AeT \max_{1 \leq i \leq k} \| \lambda_i - \hat{\lambda}(p_i) \| - \frac{x}{kT} C'_3 \delta_n \\
\overset{(b)}{\leq} x.
\]

where (a) follows from the feasibility of \( t^* \) for \((P_0)\) and the non-negativity of the elements of \( A \); and (b) follows from the conditions defining \( \mathcal{H} \) and the choice of
C_3. We see that \( \bar{t} \) is feasible for \( \hat{P} \) (for \( \omega \in \mathcal{H} \)). Let \( C'_4 = d\bar{M}T \). The revenues achieved by \( \bar{t} \) in \( \hat{P} \) can be lower bounded as follows (where \( C'_5 > 0 \) is suitably large)

\[
V_{(\hat{P})}(\bar{t}) = \sum_{i=1}^{k} p_i \cdot \hat{\lambda}(p_i) \bar{t}_i \\
\geq \sum_{i=1}^{k} p_i \cdot \hat{\lambda}(p_i) t'_i - C'_4 \eta_n \\
= \sum_{i=1}^{k} p_i \cdot \lambda_i t'_i + \sum_{i=1}^{k} p_i \cdot (\hat{\lambda}(p_i) - \lambda_i) t'_i - C'_3 \eta_n \\
\geq V_{(P_0)}(t^*) - d\bar{M} \max_{1 \leq i \leq k} \| \lambda_i - \hat{\lambda}(p_i) \| T - C'_3 \eta_n \\
\geq V_{(P_0)}(t^*) - C'_5 \max\{\delta_n, \tau_n\}.
\]

Proof of Lemma 12. The optimal vector \( \bar{p} \) for the deterministic problem is contained one of the hyper-rectangles comprising the price grid. Let \( p_j \) be the closest vector to \( \bar{p} \) in the price grid. Note that the index \( j \) depends on \( n \) but we do not make the \( n \)-dependence explicit to avoid cluttering the notation. We first show that \( p_j \in P^n_j \) with high probability. Note that \( \|p_j - \bar{p}\| \leq C'_1 / \kappa_n^{1/d} \) for some \( C'_1 > 0 \) and hence \( \|\lambda(p_j) - \lambda(\bar{p})\| \leq K C'_1 / \kappa_n^{1/d} \). We deduce that

\[
\mathbb{P}\left(p_j \notin P^n_j\right) = \mathbb{P}\left(AN(\lambda(p_j)n\Delta_nT) > n\Delta_n(x + \delta_n e)\right) \\
\leq \mathbb{P}\left(AN\left((\lambda(\bar{p}) + C'_1 K\kappa_n^{-1/d})n\Delta_nT\right) > n\Delta_n(x + \delta_n e)\right) \\
\leq \mathbb{P}\left(AN\left((\lambda(\bar{p}) + C'_1 K\kappa_n^{-1/d})n\Delta_nT\right) - A(\lambda(\bar{p}) + C'_1 K\kappa_n^{-1/d})n\Delta_nT > n\Delta_n w_n\right),
\]

where \( w_n = \delta_n e - C'_1 K T \kappa_n^{-1/d} A e \). Note that (a) is a consequence of the feasibility
of \( \hat{p} \) for the deterministic problem (in particular, \( A\lambda(\hat{p})n\Delta_nT \leq n\Delta_nx \)). Now since \( \delta_n\kappa_n^{1/d} \to \infty \), we have that \( w_n = \delta_n(e-C^*_n KT/(\delta_n\kappa_n^{1/d})Ae) \geq \delta_n/2 \) for \( n \) sufficiently large. By using Lemma 10 (where \( r_n \) and \( \varepsilon_n \) are here \( n\Delta_n \) and \( \delta_n/2 \) respectively), we deduce that the above probability is bounded above by \( C'_2/n^\eta \) for a sufficiently large \( C'_2 > 0 \). We then have

\[
\mathbb{P}(r(\hat{p}) - r(\hat{p}) > \delta_n)
\leq \mathbb{P}(r(\hat{p}) - r(\hat{p}) > \delta_n; p_j \in P^n_j, \hat{r}(p_j) > 0) + \mathbb{P}(p_j \notin P^n_j)
+ \mathbb{P}(p_j \in P^n_j, \hat{r}(p_j) = 0). \tag{C.2-5}
\]

Now under the condition that \( p^j \in P^n_j \), we have

\[
r(\hat{p}) - r(\hat{p}) = r(\hat{p}) - r(p_j) + r(p_j) - \hat{r}(p_j) + \hat{r}(p_j) - \hat{r}(\hat{p}) + \hat{r}(\hat{p}) - r(\hat{p}) \\
\leq r(\hat{p}) - r(p_j) + r(p_j) - \hat{r}(p_j) + \hat{r}(\hat{p}) - r(\hat{p}),
\]

where the inequality follows from the definition of \( \hat{p} \) given in (C.1-7). For the first term on the RHS above, note that for \( C'_3 > 0 \) suitably large

\[
|r(p_j) - r(\hat{p})| \leq |p_j \cdot \lambda(p_j) - p_j \cdot \lambda(\hat{p})| + |p_j \cdot \lambda(\hat{p}) - \hat{p} \cdot \lambda(\hat{p})|
\leq d\|p_j\|\|\lambda(p_j) - \lambda(\hat{p})\| + d\|\lambda(\hat{p})\|\|p_j - \hat{p}\|
\leq \|p_j\|K \frac{C'_1 d}{\kappa_n^{1/d}} + \|\lambda(\hat{p})\| \frac{C'_1 d}{\kappa_n^{1/d}}
\leq \frac{C'_3}{\kappa_n^{1/d}},
\]

where \( a \) follows from Cauchy-Schwarz inequality, \( b \) follows from the Lipschitz condition on \( \lambda \) and \( c \) follows from the fact that \( \|p\| \leq \bar{M} \) for all \( p \in \mathcal{D}_p \). Now,
recalling comment 2 in the preamble of Appendix C.1, we have \( r(\hat{p})T \geq m^D > 0 \)
and hence for \( n \) sufficiently large \( r(p_j) > m^D/(2T) \). By Lemma 10, we deduce that

\[
\mathbb{P}(p_j \in P^n_f, \hat{r}(p_j) = 0) \leq \mathbb{P}(p_j \cdot \tilde{\lambda}(p_j) = 0) \leq \frac{C_4}{n^\eta}
\]

Coming back to (C.2-5), since \( C_3'/\kappa_{n/1} < (1/4)\delta_n \) for \( n \) sufficiently large,

\[
\mathbb{P}(r(\hat{p}) - r(\hat{p}) > \delta_n)
\]

\[
\leq \mathbb{P}(r(p_j) - \hat{r}(p_j) > \frac{1}{2}\delta_n - \frac{C_3'}{\kappa_n^{1/4}}; p_j \in P^n_f) + \mathbb{P}(\hat{r}(\hat{p}) - r(\hat{p}) > \frac{1}{2}\delta_n; p_j \in P^n_f, \hat{r}(p_j) > 0)
\]

\[
+ \mathbb{P}(p_j \notin P^n_f) + \mathbb{P}(p_j \in P^n_f, \hat{r}(p_j) = 0)
\]

\[
\leq \mathbb{P}(r(p_j) - \hat{r}(p_j) > \frac{1}{4}\delta_n) + \mathbb{P}(\hat{p}\tilde{\lambda}(\hat{p}) - r(\hat{p}) > \frac{1}{2}\delta_n) + \frac{C_2'}{n^\eta} + \frac{C_4'}{n^\eta}
\]

By Lemma 10 the two first terms on the RHS above are bounded by \( C_5'/n^\eta \) for

some \( C_5' > 0 \) and the proof is complete. ■