Stochastic Dynamic Optimization of Consumption and the Induced Price Elasticity of Demand in Smart Grids

by

Ali Faghih
B.S., Electrical Engineering (2009)
University of Maryland, College Park

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author ....
Department of Electrical Engineering and Computer Science
May 12, 2011

Certified by. ....
Munther A. Dahleh
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ....
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

Stochastic Dynamic Optimization of Consumption and the Induced Price Elasticity of Demand in Smart Grids

by

Ali Faghih

B.S., Electrical Engineering (2009)

University of Maryland, College Park

Submitted to the Department of Electrical Engineering and Computer Science
on May 12, 2011, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

This thesis presents a mathematical model of consumer behavior in response to stochastically-varying electricity prices, and a characterization of price-elasticity of demand created by optimal utilization of storage and the flexibility to shift certain demands to periods of lower prices. The approach is based on analytical characterization of the consumer's optimal policy and the associated value function in a finite-horizon stochastic dynamic programming framework. A general model is first presented, which incorporates both load-shifting and storage, and then, the model is decoupled into two subproblems, one for load-shifting and the other for storage.

The study of optimal utilization of storage, which is performed analytically and in the presence of ramp constraints, reveals, as a particularly compelling finding, that the value function is a convex piece-wise linear function of the storage state. Moreover, it is shown that the expected monetary value of storage increases with price volatility, and that when the ramping rate is finite, the value of storage saturates quickly as the capacity increases, regardless of price volatility. Furthermore, it is shown that although the demand for electricity is often deemed to be highly inelastic, optimal utilization of local storage capacity induces a considerable amount of price elasticity of demand.

The study of the load-shifting problem is performed under both perfect and partial information about price distribution. It is shown that load-shifting induces considerable consumer savings that increase with price volatility. Furthermore, it is shown that the opportunity to optimally schedule the shiftable loads creates a considerable amount of price elasticity, even when the aggregate consumption over a long period remains insensitive to price variations.

Thesis Supervisor: Munther A. Dahleh
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

I would like to take this opportunity thank my advisor Professor Munther Dahleh whose guidance has been a great source of inspiration for me throughout the past two years. I am particularly indebted to him for his constant support and invaluable advice, and for his encouraging words during the toughest moments, which has made me realize how fortunate I am to be working under his supervision.

I would like to thank Dr. Mardavij Roozbehani for the countless hours of his time that he has spent mentoring me, providing me with insightful research ideas, and teaching me ingenious techniques for tackling research problems, without which the results in this thesis could not have been obtained. Working with Mardavij has been a great opportunity for me.

Also, I would like to thank Mesrob Ohannessian for sharing his experiences and knowledge with me throughout the past two years. Sharing an office with Mesrob has been a wonderful experience.

I am grateful to the National Science Foundation (NSF) for supporting my research through the NSF Graduate Research Fellowship.

I would like to thank Professor Edo Waks, not only for supervising my undergraduate research and boosting my motivation to pursue graduate-level research, but also for giving me invaluable advice about life. I would also like to thank Professor Christopher Davis for his priceless advice and support during my undergraduate years.

Finally, I would like to wholeheartedly thank my parents for their unconditional love and unwavering support, and for teaching me how to set my goals and providing me with the tools that I needed for achieving my goals. And my very special thanks go to my amazing sister Rose for always being there for me, being my best friend, and giving me advice with everything. I am very blessed to have my family in my life.

Contents

List of Figures

Chapter 1

Introduction

Using demand response and real-time pricing in electricity networks is considered to have numerous positive impacts on the operation of the future grid. Among these impacts is the reduction of the peak demand, which would help system operators reduce the maximum required capacity and the high costs of capacity expansion. The reduction in the peak demand would result in mitigation or even elimination of the need for high carbon emission power plants that are brought in to the grid for only a very short period of time per year just to meet the annual peak demand. Another notable impact is minimization of the reserve capacity needed for meeting the demand in the event of contingencies. Hence, the study of demand response in the presence of time-varying electricity prices should be of utmost importance to system designers and system operators.

Characterization of demand response, however, requires the study of the technologies that are directly related to, or highly influence, the qualitative and quantitative characteristics of demand response. The rapidly growing demand for electricity and the urge to reduce green-house emissions necessitate incorporation of a large number of renewable energy sources and highly efficient sustainable technologies in the future grid. Toward this end, a substantial number of consumers are expected to adopt real-time demand response technologies to minimize the expected costs. However, the anticipated large-scale integration of sustainable technologies will considerably add to the uncertainty faced by the system operators. Hence, there is a need for development of econometric models of consumer behavior that enable system operators to characterize the responsiveness of demand to stochastically-varying electricity prices in the presence of these technologies.

Availability of such econometric models to the system operators is of considerable importance, particularly for maintaining the stability of the system. For instance, it was shown in [21] that in real-time priced power grids with information asymmetry, the stability and robustness of the system to disturbances are greatly affected by the consumers' valuation of electricity, whereas in electricity markets in which the central market operator has full information about the consumers' valuation of electricity, the situation is quite different. In these markets, if the demand becomes even moderately price responsive, the magnitude and frequency of price spikes would be substantially mitigated, and the average spot price of electrical energy would decrease [14], [6], [9]; on the other hand, low price-elasticity of demand could cause large price spikes in spot electricity markets [14], and pave the way for generation companies to exercise market power [7].

Nevertheless, the demand for electricity has often been considered to be highly inelastic. In this regard, [23] and [11] have asserted that demand decreases in response to a short-term price increase only to a relatively small extent. Hence, having a modern and more efficient power grid calls for integration of real-time demand response technologies that would considerably increase the responsiveness of demand to stochastically-varying prices.

Two such technologies are energy storage and load-shifting, which are the focus of this thesis. Characterization of the limitations and implications of optimal management of storage and load-shifting, and how they affect the price elasticity of demand would be of importance to various entities, from consumers to system designers and system operators. This thesis seeks to provide such characterization by presenting a model for optimal utilization of storage capacity and load-shifting in response to stochastically-varying electricity prices, and characterizing the price-elasticity of demand induced by adoption of this model. In this thesis, the load-shifting and the storage management problem are formulated in a finite-horizon stochastic dynamic programming framework, and analytical expressions are given for the optimal policy and the corresponding value function, particularly, revealing that the value function in the storage problem is indeed a convex piece-wise linear function of the storage state. An important feature of the model formulated and solved in this thesis is that it takes the physical ramp constraints of storage into account. The physical ramp constraints, which limit the storage system's ability to move between different operating levels over short periods of time, make characterization of analytical solutions particularly difficult. Consumer savings induced by optimal load-shifting and the monetary value of storage for the consumer, as well as the price elasticity of demand induced by optimal load-shifting and storage management, are all analyzed within the same mathematical framework.

Shiftable loads are situated in a variety of forms for various consumers. Across the full spectrum of residential, commercial and industrial consumption, at any given time, a considerable portion of the generated power is supplied to shiftable loads that are deferrable for a few minutes, or possibly hours, at little or no cost [19]. Examples include electric vehicle charging, heating, ventilation, air conditioning, refrigeration, agricultural pumping, laundry and dishwashing. Some loads such as refrigeration, air-conditioning and heating can also be viewed as thermal storage via pre-cooling or pre-heating. Electric vehicle charging can be viewed as both electrical storage and shiftable load. Battery energy storage systems and hydro-electric storage systems are two other examples of electrical energy storage systems that could be optimally managed.

Related quantitative frameworks generally appear in the literature that address the consumer energy management problem, and are mostly based on stochastic dynamic programming. Some earlier works such as [10],[5] have laid the groundwork and introduced the general concepts, and some recent studies [17], [19], have delved deeper into the concepts, and obtained results that are more related to this work.

Livengood and Larson's work [17] proposes the design of a software energy management system for the typical small consumer of electricity. This software consists of a set of algorithms that use stochastic dynamic programming for optimal scheduling and management of the consumer's electricity consumption, storage, and selling back to the grid in the face of uncertain electricity prices and weather conditions. Although the idea behind the formulation of their model is very similar to that of this thesis, their approach is quite different. In contrast to their model, the model presented in this thesis is more abstract, which allows us to derive analytical expressions for the optimal solution of the underlying dynamic programming problem. The results obtained through the analytical approach of this thesis give an abstract description of a complicated behavioral model, which not only provides a model for optimal management of consumption, but also allows us to develop simplified models that effectively highlight the essential structural features of consumer behavior from the system operator's point of view.

Regarding load-shifting, a closely related work by Papavasiliou and Oren [19] proposes a direct coupling of renewable generation with shiftable loads to mitigate the imbalances caused by the unpredictable and uncontrollable fluctuation of renewable energy supply. Their approach too, is based on stochastic dynamic programming. In particular, the results presented in this thesis on the affine structure of the value function associated with the optimal load-shifting problem were partially obtained in [19].

The literature covering the topic of energy storage is extensive. Different variations of the storage problem have been addressed in various contexts and from different aspects, to serve different objectives. Formulating optimization models for scheduling electricity storage devices has been the topic of several previous works such as [1] and [18]. Lee and Chen [15, 16] study industrial customers with time-of-use rates and determine optimal contracts and optimal sizes of battery storage systems for such consumers in a dynamic programming framework. Moreover, Bannister and Kaye in [3] focus their study on optimizing the operation of a single storage connected to a general linear memoryless system in the presence of ramp constraints. Their approach is based on linear optimization and deterministic dynamic programming. Although their model is somewhat similar to the model in this thesis in the sense that it deals with optimal utilization of storage in the presence of ramp constraints, the deterministic nature of their approach makes the mathematical framework of their model, and hence their conclusions, quite different from those of this thesis. The economic benefits of electricity storage to the end consumer have also been reported in previous studies such as [2]. In contrast, several other works such as [13], [8], [22], [4], and [12] have studied the idea of employing energy storage for efficient integration of renewable sources.

In particular, in [12], the optimal storage investment problem for efficient integration of renewable sources is studied in an infinite-horizon stochastic dynamic programming framework. The storage management problem presented in [12] is formulated and solved from the point of view of a renewable generation owner who wants to fulfill her on-site (local) demand using her renewable generator. At each time step, if the on-site demand is lower than the renewable generation, the generation owner uses a storage device to store any generation that is in excess of her on-site demand; on the other hand, if the renewable generation is lower than the on-site demand, any excess demand that is not satisfied from renewable generation is fulfilled by purchasing from other generators connected to the main grid at prices which are stochastic and revealed right before consumption. Hence, the purpose of the storage model in [12] is to store local renewable generation only and not store energy from the main grid (selling back to the grid is not allowed either). Particularly, the focus of [12] is on the case when the renewable generation owner is given either a low price or a high price each with known probabilities, and then, from various perspectives, optimal sizing of energy storage is characterized in the presence of this price distribution. Although the underlying idea of optimally utilizing limited storage capacity in response to stochastically varying prices and characterizing the value of storage capacity to the consumer presented in [12] is similar to that of this thesis, there are fundamental differences between the objectives, formulation, approach, and hence, the results of the two works, which put the contributions of the two studies in quite different frameworks. First, in contrast to [12], the model in this thesis allows selling stored energy back to the main grid, which creates a considerable difference in the value of storage compared to the one studied in [12]; the model in this thesis, however, assumes that the consumer only interacts with the main grid and does not have access to on-site renewable generation. Second, the model in this thesis does not assume a specific distribution on prices, and instead, provides analytical expressions for the consumer's optimal threshold policy that could be applied to any price distribution, which allows this thesis to compare and contrast the value of storage under different price distributions. In contrast, in [12], the presented policy is not based on deriving thresholds, and is optimal only for pricing schemes that have at most two price levels. This approach, in turn, has allowed them to derive very specific bounds for optimal sizing of storage in the presence of such two-level pricing schemes; their findings reveal that any investment in storage is profitable only if the ratio of the amortized capital cost of storage to the higher price-level of energy is less than 1/4. Another very important difference between the two studies is that this thesis analytically characterizes the effect of ramp constraints on the structure of the value function, and illustrates how the ratio of the storage capacity to the physical ramp constraints affects the value of storage; though, such characterization of the effects of ramp constraints has not been a topic of interest in [12]. Hence, considering all the aforementioned differences between the value functions of the two problems, the upper-bound derived on the cost of storage in [12] is not applicable to the storage problem presented in this thesis. In summary, the possibility of selling back to grid, absence of local renewable generation, allowing arbitrary price distributions, and the piece-wise linearity of the value function resulted from ramp constraints, which are all incorporated into or highlighted by the model presented in this thesis, put the two works in different frameworks, both from the mathematical and the qualitative point of view.

Another closely related work is presented in [4], in which Bitar et al. study the impact of energy storage capabilities on revenue of a wind power producer. While the main focus of their study, and hence their results and conclusions are quite different from those of this thesis, their formulation of the underlying stochastic optimal control problem is fairly similar to the one presented in this thesis.

The existing literature on price-elasticity of the demand for electricity are mostly based on empirical evidence and qualitative reasoning, see, for instance, [14], [23], and [11]. One of the goals of this thesis is to address price-elasticity in a quantitative framework, and illustrate that a considerable increase in the price-elasticity of demand could be obtained through optimal utilization of storage capacity and optimal scheduling of shiftable loads.

In the following chapters, a mathematical model of consumption is first formulated, and then this model is decoupled into two subproblems: the storage problem and the load-shifting problem. Chapters 3 and 4 study each of these two sub-problems separately.

The contributions of this thesis regarding the storage problem can be summarized as follows:

- A behavioral model of the consumer is proposed, which is based on an optimal policy for managing storage in the presence of ramp constraints. The solution of the underlying stochastic dynamic program is analytically characterized, and as a particularly interesting finding of this thesis, it is shown that at each instant of time, the value function is a convex piece-wise linear function of the storage state.

- It is shown that the expected monetary value of storage capacity increases with price volatility. Moreover, it is shown that when the ramp constraint is finite, the value of storage saturates quickly as the capacity increases, regardless of price volatility

- Finally, a behavioral model that represents the aggregate response of a large number of consumers is provided, and an expression for the price-elasticity of the aggregate demand is presented, which highlights the fact that an individual consumer's response to a price signal is dependent on both the price and the internal state of the consumer. It is shown that optimal scheduling of local storage capacity induces a considerable amount of price elasticity of demand.

The contributions of this thesis regarding the load-shifting problem can be summarized as follows:

- A behavioral model for load-shifting is characterized based on analytical expressions of the optimal threshold policy, under both perfect and partial information about price distribution.
- The model is used to show that consumer's expected savings from optimal load-shifting is an increasing function of price volatility. The relation between expected savings and price volatility is examined through analytical bounds for simple distributions and through simulations when analytical bounds could not be established.
- Finally, it is shown that optimal load-shifting can create a considerable amount of short-term price elasticity, even when the cumulative consumption over a long period remains insensitive to price variations.

Chapter 2

Formulation of the Dynamic Consumption Management Model

In this chapter, a model for dynamic optimization of consumption in the presence of stochastically-varying electricity prices, load-shifting, and storage capacity is presented. This model is formulated based on the principles of stochastic dynamic programming, and it provides a behavioral model of the consumer, which will be exploited in the following chapters to characterize the intertemporal utility of consumption induced by load-shifting and storage.

An important feature of this model is that it takes the physical ramp constraints of storage into account. Moreover, a parametric upper bound is imposed on the amount of storage available to the consumer, and the effect of positive storage capacity on the net consumption is investigated from an Input/Output behavioral point of view. Negative storage or backlogging is allowed for certain types of flexible loads that are shiftable in time, while a deadline is imposed for fulfilling the shifted loads.

The notations and model components are defined in the first section, and the optimization-based model is presented in the second section.

2.1 Model Components

The consumer's energy management problem is formulated as an inventory control problem over a finite time-horizon. The following subsections define and characterize the components that will be used in the next section to formulate the energy management problem.

Notations

The set of positive real numbers (integers) is denoted by $\mathbb{R}_+$ ($\mathbb{Z}_+$), non-negative real numbers (integers) by $\overline{\mathbb{R}}_+$ ($\overline{\mathbb{Z}}_+$), and similarly for their negative and non-positive counterparts.

Demand

The consumer's total demand at time $k \in \{0, \cdots, N\}$ is denoted by d_k. It is assumed that the demand consists of two components:

$$d_k = d_k^f + d_k^s$$

where d_k^f is the firm component and d_k^s is the shiftable component.

At time $k = 0$, both d_k^s and d_k^f are perfectly known to the consumer for all periods $k \in \{0, ..., N\}$. The shiftable demand d_k^s can be satisfied at any time $t \in \{k, ..., N\}$, whereas the fixed demand d_k^f must be satisfied at time k.

Note: Both d_k^s and d_k^f are assumed to be inelastic. Hence, the results obtained on price-elasticity of demand in chapters 3 and 4 of this thesis characterize the price-elasticity of demand induced solely by storage and shifting, respectively.

The Decisions

The decision set of the consumer is characterized by a triplet

$$(u_k, v_k^{\text{in}}, v_k^{\text{out}}) \in [0, \overline{u}) \times [0, \overline{v}^{\text{in}}] \times [0, \overline{v}^{\text{out}}] \tag{2.1}$$

where, u_k is the amount of electricity that, at time k, the consumer allocates to fulfilling some or all of the shiftable demands, and v_k^{in} and v_k^{out} are, respectively, the amount of electricity that the consumer stores in, or withdraws from the storage. The corresponding upper bounds ($\overline{v}^{\text{in}}$ and $\overline{v}^{\text{out}}$) represent the physical ramp constraints on storage.

A simplified model can also be obtained, in which the decision set is characterized by a pair

$$(u_k, v_k) \in [0, \overline{u}) \times [-\overline{v}^{\text{out}}, \overline{v}^{\text{in}}] \tag{2.2}$$

where, u_k is defined as before and v_k can represent both storing energy (when $v_k \geq 0$) and withdrawing energy from the storage (when $v_k \leq 0$).

The net consumption (total purchase or sell-back to the grid) is then given by

$$y_k = v_k^{\text{in}} - v_k^{\text{out}} + u_k + d_k^f \tag{2.3}$$

or, alternatively, by

$$y_k = v_k + u_k + d_k^f \tag{2.4}$$

In either case, it is assumed that y_k is constrained as:

$$y_k \in [-\underline{y}, \overline{y}], \qquad \underline{y}, \overline{y} \in [0, \infty) \tag{2.5}$$

Hence, $y_k < 0$ is associated with selling electricity back to the grid, and $\underline{y} = 0$ corresponds to the situation where selling electricity back to the grid is not allowed.

The Price

The price process λ_k is assumed to be an exogenous Markovian process driven by an independently distributed random process $\mathbf{w}_k$ according to

$$\lambda_{k+1} = g\left(\lambda_k, w_k\right)$$

where the function g_k and the distributions of $\mathbf{w}_k$ are assumed known for each k. The support of the price distribution is assumed to be non-negative, and is denoted by λ_{min} and λ_{max} (i.e. the prices are distributed between λ_{min} and λ_{max}, such that $0 \leq \lambda_{min} \leq \lambda_{max}$). It is assumed that at the beginning of each discrete time interval $[k, k+1]$, the random variable λ_k is materialized and revealed to the consumer. A specific scenario where this model is readily applicable is where the distributions of the prices for the next 24 hours are computed in the day-ahead market and made available to the consumer. In this case, one may choose $g\left(\lambda_k, w_k\right) = w_k$, where the distribution of w_k is known.

It is also assumed that the feed-in and usage tariffs are the same. That is, for each pricing interval, λ_k is the price per unit for both consumption (corresponding to $y_k \geq 0$) and production (i.e. negative consumption, corresponding to $y_k \leq 0$), and there are no transaction costs.

The States

It is assumed that the consumer holds an energy inventory characterized by a pair

$$(x_k, s_k) \in (-\infty, 0] \times [0, \overline{s}] \tag{2.6}$$

where x_k represents the amount of backlogged/shifted demand, and s_k represents the energy stored in the local storage.

Note that we impose a deadline on backlog by constraining $x_N = 0$. Also, the parameter $\overline{s}$ is the physical upper bound on the amount of storage available to the consumer (hence, $\overline{s} = 0$ corresponds to the case of no storage capacity).

The states x_k and s_k evolve according to:

$$x_{k+1} = x_k + u_k - d_k^s \tag{2.7}$$

$$s_{k+1} = \beta s_k + \eta^{\text{in}} v_k^{\text{in}} - \eta^{\text{out}} v_k^{\text{out}} \tag{2.8}$$

where, u_k, v_k^{in}, and v_k^{out} are defined as in (2.1), $\beta \leq 1$ is the decay factor, $\eta^{\text{in}} \leq 1$ and $\eta^{\text{out}} \geq 1$ are charging and discharging efficiency factors. Note that although in this model the efficiency factors and the ramp rates are assumed to be constants, they might, in general, be complicated functions of the operating point (i.e., the storage level) in certain practical scenarios.

The idealized model of the dynamics of storage can be written as:

$$s_{k+1} = s_k + v_k, \qquad v_k \in [-\overline{v}_{\text{out}}, \overline{v}_{\text{in}}] \tag{2.9}$$

which corresponds to $\beta = 1$, $\eta^{\text{in}} = 1$, and $\eta^{\text{out}} = 1$.

Disutility and Penalty

It is assumed that in general, there is a disutility associated with backlogging the demand. This disutility is characterized via a cost function $p_k(\cdot)$ which essentially represents, in an abstract sense, the inconvenience that the consumer experiences for fulfilling some of her shiftable demands at a future time.

Furthermore, there is a penalty associated with storage via cost function $h_k(\cdot)$, which characterizes certain costs that the consumer may incur for access to and/or use of storage.

2.2 The Optimization-Based Model

Using the components defined in the previous section, the consumer's energy management problem can be formulated as a finite-horizon stochastic dynamic programming problem as follows:

$$
\begin{aligned}
\min\quad & \mathbf{E}\left[\sum_{k=0}^{N} p_k(x_k) + h_k(s_k) + \lambda_k y_k\right] \\
\text{s.t.}\quad & x_{k+1} = x_k + u_k - d_k^s, \qquad x_N = 0 \\
& s_{k+1} = \beta s_k + \eta^{\text{in}} v_k^{\text{in}} - \eta^{\text{out}} v_k^{\text{out}} \\
& \lambda_{k+1} = g(\lambda_k, w_k) \\
& y_k = u_k + v_k^{\text{in}} - v_k^{\text{out}} + d_k^f \\
& (u_k, v_k^{\text{in}}, v_k^{\text{out}}) \in [0, \overline{u}) \times [0, \overline{v}^{\text{in}}] \times [0, \overline{v}^{\text{out}}] \\
& (x_k, s_k) \in (-\infty, 0] \times [0, \overline{s}] \\
& y_k \in [-\underline{y}, \overline{y}], \qquad \underline{y}, \overline{y} \in [0, \infty)
\end{aligned}
\tag{2.10}
$$

This optimization problem can be decoupled into two subproblems. Consider the following optimization problems:

1. *The storage problem*:

$$
\begin{aligned}
\min\quad & \mathbf{E}\left[\sum_{k=0}^{N} h_k(s_k) + \lambda_k v_k\right] \\
\text{s.t.}\quad & s_{k+1} = s_k + v_k \\
& \lambda_{k+1} = g(\lambda_k, w_k) \\
& s_k \in [0, \overline{s}] \\
& v_k \in [-\overline{v}^{\text{out}}, \overline{v}^{\text{in}}]
\end{aligned}
\tag{2.11}
$$

2. *The load-shifting problem:*

$$\begin{aligned}
\min \quad & \mathbf{E}\left[\sum\nolimits_{k=0}^{N} p_k(x_k) + \lambda_k u_k\right] \qquad (2.12)\\
\text{s.t.} \quad & x_{k+1} = x_k + u_k - d_k^s, \qquad x_N = 0\\
& \lambda_{k+1} = g(\lambda_k, w_k)\\
& x_k \in (-\infty, 0]\\
& u_k \in [0, \overline{u})
\end{aligned}$$

Proposition. *Let γ_c^*, γ_x^*, and γ_s^* be, respectively, the optimal solutions to the consumer's optimization problem (2.10), the load-shifting problem (2.12), and the storage problem (2.11). Let $\bar{\gamma}_f$ be the expected cost of the firm demands:*

$$\bar{\gamma}_f = \mathbf{E}\left[\sum\nolimits_{k=0}^{N} \lambda_k d_k^f\right]$$

Then, for sufficiently large $\overline{u}$:

$$\gamma_c^* \geq \gamma_x^* + \gamma_s^* + \bar{\gamma}_f$$

Furthermore, suppose that both $\overline{y}$ and $\overline{u}$ are sufficiently large, and that the storage is lossless, i.e., $\beta = 1, \eta^{in} = 1$, and $\eta^{out} = 1$. Then

$$\gamma_c^* = \gamma_x^* + \gamma_s^* + \bar{\gamma}_f$$

The implication of the above proposition is that when the feeder limits ($\overline{y}$ and $\overline{u}$) are sufficiently large, the storage is ideal, and the feed-in and usage tariffs are the same, then, the consumer is indifferent to satisfying the demand by withdrawing from the grid, or from the storage. Hence, not only the problems of storage and load-shifting can be solved separately, but also, the demand profile d_k^f becomes irrelevant in decision-making. Since the goal of this thesis is to develop simplified models that effectively highlight the essential structural features of consumer behavior, the idealized model of storage will be adopted in this study.

Remark 1. *Although in the formulation of the storage problem (2.11) it was assumed that the storage is ideal, a storage penalty component* $h_k(s_k)$ *has been incorporated into the cost function, which, in an abstract sense, could be used as a proxy to account for the losses due to non-ideal storage.*

Remark 2. *The storage problem is formulated and solved only for the finite-horizon case. Considering that it is somewhat unrealistic to assume a certain distribution for prices far in the future, the infinite horizon case may not yield a realistic model of consumer behavior. However, the finite horizon problem will be solved in such a way that the ongoing process of storage after the end of stage* N *is taken into account. In other words, the finite time horizon of the storage problem is not treated as a deadline; rather, a value of* $\overline{\lambda}$ *is assigend to each unit of energy left in storage by the end of the time horizon (i.e. at stage* N*), based on the idea that the consumer will be able to use the energy in her storage in the period that follows the current time-horizon.*

Chapter 3

The Storage Problem

This chapter studies the storage problem defined in (2.11). An analytical characterization of the optimal policy and the associated value function is presented in the first section, and then, in the next two sections, these analytical results are used to evaluate the expected monetary value of storage and the induced price-elasticity of demand.

3.1 The Optimal Policy

In this section, the consumer's policy for optimal utilization of storage capacity is characterized based on principles of stochastic dynamic programming and optimal control. This allows us to develop a mathematical model for the intertemporal utility of consumption, induced by storage capacity. It is assumed that the price distribution is perfectly known.

In particular, in this section, it is shown that the value function is a convex piecewise linear function of the storage state. This result is quite compelling, but non-trivial.

Theorem 1. *Assume that the price distribution has finite support. Consider the* *storage problem (2.11) with*

$$\overline{v}^{in} = \overline{v}^{out} = \overline{v}$$

where $\overline{s} = n\overline{v},\ n \in \mathbb{Z}_+$, *and the penalty functions* $h_k : [0,\infty) \to [0,\infty),\ k = 0,..,N$ *are piecewise linear convex functions of the form:*

$$h_k(s) = h_{i,k}s + c_{i,k},\ \ s \in [i\overline{v},(i+1)\overline{v})\,, i \in \overline{\mathbb{Z}}_+ \tag{3.1}$$

where

$$0 \le h_{i,k} \le h_{i+1,k}, \qquad \forall k, i$$

Then

(i) The value function is a convex piecewise-linear function of the form:

$$V_k(s) = -t_{i,k}s + e_{i,k},\ \ s \in [i\overline{v},(i+1)\overline{v})\,, i \in \overline{\mathbb{Z}}_+ \tag{3.2}$$

where

$$t_{i+1,k} \le t_{i,k}, \qquad \forall k, i$$

(ii) The optimal policy is a threshold policy characterized by:

if $0 \le s_k < \overline{v}$ *(i.e.* $i = 0$*), then*

$$v_k^* = \begin{cases} -s_k & \text{if } \ t_{0,k+1} < \lambda_k \\ \overline{v} - s_k & \text{if } \ t_{1,k+1} < \lambda_k \le t_{0,k+1} \\ \overline{v} & \text{if } \ \lambda_k \le t_{1,k+1} \end{cases}$$

if $s_k \ge \overline{v}$*, such that* $s_k \in [i\overline{v},(i+1)\overline{v})$*,* $i \in 1,2,...,n-1$*, then*

$$v_k^* = \begin{cases} -\overline{v} & \text{if } \ t_{i-1,k+1} < \lambda_k \\ i\overline{v} - s_k & \text{if } \ t_{i,k+1} < \lambda_k \le t_{i-1,k+1} \\ (i+1)\overline{v} - s_k & \text{if } \ t_{i+1,k+1} < \lambda_k \le t_{i,k+1} \\ \overline{v} & \text{if } \ \lambda_k \le t_{i+1,k+1} \end{cases}$$

and the thresholds are given by the recursive equations:

$$t_{i,N} = \bar{\lambda}, \qquad i \in 0, 1, 2, ..., n-1$$

$$t_{i,N} = -h_{i,N}, \qquad i \geq n$$

for $k < N$ *and* $i \in \mathbb{Z}_+$:

$$t_{0,k} = t_{1,k+1} F_\Lambda \left(t_{1,k+1}\right) - h_{0,k} + \sum_{t_{1,k+1} < \theta \leq \lambda_{max}} \theta P_\Lambda(\theta)$$

$$t_{i,k} = t_{i-1,k+1} \left(1 - F_\Lambda \left(t_{i-1,k+1}\right)\right) + t_{i+1,k+1} F_\Lambda \left(t_{i+1,k+1}\right) - h_{i,k} + \sum_{t_{i+1,k+1} < \theta \leq t_{i-1,k+1}} \theta P_\Lambda(\theta)$$

where, $P_\Lambda(\cdot)$ *and* $F_\Lambda(\cdot)$ *denote the probability mass function (pmf) and the cumulative distribution function (cdf) of prices, respectively. Note that the above results are expressed in the form of discrete probability distributions, but they extend naturally to continuous distributions.*

Proof. This proof proceeds by induction. Let us for the moment assume that the value function, $V_k(\cdot) \equiv \mathbf{E}\left[J_k\left(\cdot\right)\right]$, has the form defined in (3.2).
From the dynamic programming algorithm, for $k < N$, we have

$$J_k\left(s_k\right) = h_k\left(s_k\right) + \min_{v_k \in [max(-s_k, -\overline{v}), \overline{v}]} \left\{\lambda_k v_k + \mathbf{E}\left[J_{k+1}\left(s_k + v_k\right)\right]\right\} \tag{3.3}$$

where the penalty functions $h_k(s_k)$ are as defined in (3.1).
Then, using the general form in (3.2) for $\mathbf{E}\left[J_k\left(s_k\right)\right]$ and the state evolution equation in (2.6), and by applying the induction step to (3.3), for $i\overline{v} \leq s_k + v_k < (i+1)\overline{v}$, $i \in \overline{\mathbb{Z}}_+$, we obtain:

$$J_k(s_k) \;=\; h_{i,k} s_k \,+\, c_{i,k} \,+ \min_{v_k \in [max(-s_k, -\overline{v}), \overline{v}]} \left\{\lambda_k v_k \,-\, t_{i,k+1}(s_k \,+\, v_k) \,+\, e_{i,k+1}\right\} \tag{3.4}$$

Hence, if $0 \leq s_k < \overline{v}$ (i.e. $i = 0$), then

$$v_k^* = \begin{cases} -s_k & \text{if } \; t_{0,k+1} < \lambda_k \\ \overline{v} - s_k & \text{if } \; t_{1,k+1} < \lambda_k \leq t_{0,k+1} \\ \overline{v} & \text{if } \; \lambda_k \quad \leq t_{1,k+1} \end{cases}$$

and if $s_k \geq \overline{v}$, such that $s_k \in [i\overline{v}, (i+1)\overline{v})$, $i \in 1, 2, ..., n-1$, then

$$v_k^* = \begin{cases} -\overline{v} & \text{if } \; t_{i-1,k+1} < \lambda_k \\ i\overline{v} - s_k & \text{if } \; t_{i,k+1} \quad < \lambda_k \leq t_{i-1,k+1} \\ (i+1)\overline{v} - s_k & \text{if } \; t_{i+1,k+1} < \lambda_k \leq t_{i,k+1} \\ \overline{v} & \text{if } \; \lambda_k \quad \leq t_{i+1,k+1} \end{cases}$$

Therefore, if $0 \leq s_k < \overline{v}$:

$$J_k(s_k) = \begin{cases} (h_{0,k} - \lambda_k)s_k + e_{0,k+1} + c_{0,k} & \text{if } \; t_{0,k+1} < \lambda_k \\ (\lambda_k - t_{1,k+1})\overline{v} + (h_{0,k} - \lambda_k)s_k + e_{1,k+1} + c_{0,k} & \text{if } \; t_{1,k+1} < \lambda_k \leq t_{0,k+1} \\ (\lambda_k - t_{1,k+1})\overline{v} + (h_{0,k} - t_{1,k+1})s_k + e_{1,k+1} + c_{0,k} & \text{if } \; \lambda_k \quad \leq t_{1,k+1} \end{cases}$$

and if $s_k \geq \overline{v}$, such that $s_k \in [i\overline{v}, (i+1)\overline{v})$, then

$$J_k(s_k) = \begin{cases} -(\lambda_k - t_{i-1,k+1})\overline{v} + (h_{i,k} - t_{i-1,k+1})s_k + e_{i-1,k+1} + c_{i,k} & \text{if } \; t_{i-1,k+1} < \lambda_k \\ (\lambda_k - t_{i,k+1})i\overline{v} + (h_{i,k} - \lambda_k)s_k + e_{i,k+1} + c_{i,k} & \text{if } \; t_{i,k+1} \quad < \lambda_k \leq t_{i-1,k+1} \\ (\lambda_k - t_{i+1,k+1})(i+1)\overline{v} + (h_{i,k} - \lambda_k)s_k + e_{i+1,k+1} + c_{i,k} & \text{if } \; t_{i+1,k+1} < \lambda_k \leq t_{i,k+1} \\ (\lambda_k - t_{i+1,k+1})\overline{v} + (h_{i,k} - t_{i+1,k+1})s_k + e_{i+1,k+1} + c_{i,k} & \text{if } \; \lambda_k \quad \leq t_{i+1,k+1} \end{cases}$$

Let us recall that for a function

$$f(x) = \begin{cases} f_1(r)x + g_1(r) & \text{if } \; r \leq \alpha \\ f_2(r)x + g_2(r) & \text{if } \; r > \alpha \end{cases}$$

we have

$$\begin{aligned}
\mathbf{E}[f(x)] &= \mathbf{E}[f(x)\,|r \le \alpha]\,\mathbf{P}(r \le \alpha) + \mathbf{E}[f(x)\,|r > \alpha]\,\mathbf{P}(r > \alpha) \\
&= \mathbf{E}[f_1(c)\,x + g_1(r)\,|r \le \alpha]\,\mathbf{P}(r \le \alpha) + \mathbf{E}[f_2(r)\,x + g_2(c)\,|r > \alpha]\,\mathbf{P}(r > \alpha) \\
&= x\left[\mathbf{E}[f_1(r)\,|r \le \alpha]\,\mathbf{P}(r \le \alpha) + \mathbf{E}[f_2(r)\,|r > \alpha]\,\mathbf{P}(r > \alpha)\right] \\
&\quad + \mathbf{E}[g_1(r)\,|r \le \alpha]\,\mathbf{P}(r \le \alpha) + \mathbf{E}[g_2(r)\,|r > \alpha]\,\mathbf{P}(r > \alpha) \qquad (3.5)
\end{aligned}$$

where,

$$\mathbf{E}[f(r)\,|r \le \alpha]\,\mathbf{P}(r \le \alpha) = \sum_{\theta=r_{\min}}^{\alpha} f(r)\,P_R(\theta) \qquad (3.6)$$

Now, let us apply the method described in (3.5) and (3.6) to the equations derived above for $J_k(s_k)$, and compute $\mathbf{E}[J_k(s_k)]$ for $k < N$, which leads to the following results:

if $0 \le s_k < \overline{v}$, then

$$\mathbf{E}[J_k(s_k)] = e_{0,k} - s_k[t_{1,k+1}F_\Lambda(t_{1,k+1}) - h_{0,k} + \sum_{t_{1,k+1}<\theta\le\lambda_{max}} \theta P_\Lambda(\theta)]$$

if $s_k \ge \overline{v}$, such that $s_k \in [i\overline{v}, (i+1)\overline{v})$, $i \in \mathbb{Z}_+$, then

$$\mathbf{E}[J_k(s_k)] = e_{i,k} - s_k[t_{i-1,k+1}(1 - F_\Lambda(t_{i-1,k+1})) + t_{i+1,k+1}F_\Lambda(t_{i+1,k+1}) - h_{i,k} + \sum_{t_{i+1,k+1}<\theta\le t_{i-1,k+1}} \theta P_\Lambda(\theta)]$$

where $e_{i,k}$ denotes the sum of the terms that have not been multiplied by s_k. Hence, the thresholds for $k < N$ and $i \in \mathbb{Z}_+$ are:

$$\begin{aligned}
t_{0,k} &= t_{1,k+1}F_\Lambda(t_{1,k+1}) - h_{0,k} + \sum_{t_{1,k+1}<\theta\le\lambda_{max}} \theta P_\Lambda(\theta) \\
t_{i,k} &= t_{i-1,k+1}(1 - F_\Lambda(t_{i-1,k+1})) + t_{i+1,k+1}F_\Lambda(t_{i+1,k+1}) - h_{i,k} + \sum_{t_{i+1,k+1}<\theta\le t_{i-1,k+1}} \theta P_\Lambda(\theta)
\end{aligned}$$

The next step is to verify, using induction, that the thresholds at each stage (i.e. $t_{i,k}$) are a non-increasing function of i. In order to do so, considering that $t_{0,k-1}$ has a different general form than $t_{i,k-1}$ for $i > 0$, we first need to show that $t_{1,k-1} \leq t_{0,k-1}$ assuming that $t_{i,k}$ was a non-increasing function of i. Then, we need to use induction to show that $t_{i+1,k-1} \leq t_{i,k-1}$ for $i \in \mathbb{Z}_+$, assuming that $t_{i,k}$ was a non-increasing function of i. Hence, we start by showing that $t_{i+1,N} \leq t_{i,N}$, and then, based on the assumption that $t_{i,k}$ was a non-increasing function of i, show that $t_{i+1,k-1} \leq t_{i,k-1}$ for $i \in \mathbb{Z}_+$.

As the first step, let us verify that $t_{0,k-1}, \geq t_{1,k-1}$, assuming that $t_{i,k}$ was a non-increasing function of i. Hence, we want to show that

$$\begin{aligned} t_{1,k}\left(F_\Lambda\left(t_{1,k}\right)\right) - h_{0,k} + \sum_{t_{1,k}<\theta\leq\lambda_{max}} \theta P_\Lambda\left(\theta\right) \quad &\geq \quad t_{0,k}\left(1 - F_\Lambda\left(t_{0,k}\right)\right) + t_{2,k}\left(F_\Lambda\left(t_{2,k}\right)\right) - \\ & \qquad h_{1,k} + \sum_{t_{2,k}<\theta\leq t_{0,k}} \theta P_\Lambda\left(\theta\right) \end{aligned}$$

Knowing that $h_{1,k} \geq h_{0,k}$, we can remove $-h_{1,k}$ and $-h_{0,k}$ from both sides; then, by taking the negative terms from each side to the other side to make all the terms positive on both sides, and by writing the cumulative distribution functions as summations, the above can be rewritten as:

$$\begin{aligned} t_{0,k} \sum_{\lambda_{min}\leq\theta\leq t_{0,k}} P_\Lambda\left(\theta\right) + \sum_{\lambda_{min}\leq\theta\leq t_{1,k}} P_\Lambda\left(\theta\right) t_{1,k} \\ + \sum_{t_{1,k}<\theta\leq\lambda_{max}} \theta P_\Lambda\left(\theta\right) \geq t_{0,k} + \sum_{\lambda_{min}\leq\theta\leq t_{2,k}} P_\Lambda\left(\theta\right) t_{2,k} + \sum_{t_{2,k}<\theta\leq t_{0,k}} \theta P_\Lambda\left(\theta\right) \end{aligned}$$

By breaking each summation into disjoint intervals, and factoring all the terms in the same interval and merging them into one summation, the above can be rewritten as:

$$\begin{aligned} t_{0,k} \leq & \sum_{\lambda_{min}\leq\theta\leq t_{2,k}} P_\Lambda\left(\theta\right)\left(t_{0,k} + t_{1,k} - t_{2,k}\right) + \sum_{t_{2,k}<\theta\leq t_{1,k}} P_\Lambda\left(\theta\right)\left(t_{0,k} + t_{1,k} - \theta\right) + \\ & \sum_{t_{1,k}<\theta\leq t_{0,k}} P_\Lambda\left(\theta\right) t_{0,k} + \sum_{t_{0,k}<\theta\leq\lambda_{max}} \theta P_\Lambda\left(\theta\right)) \end{aligned}$$

We can see by inspection that the above inequality is true. We can see that in the equation on the right hand side (RHS) of the inequality shown above, all the terms that have been multiplied by $P_\Lambda\left(\theta\right)$ inside the summations are greater than or equal to $t_{0,k}$; we also know that $\sum_{\theta=\lambda_{min}}^{\lambda_{max}} P_\Lambda(\theta) = 1$. Hence, we can clearly see that the RHS equation in the inequality shown above is always greater than or equal to $t_{0,k}$.

Now we can use induction to show that $t_{i+1,k-1} \leq t_{i,k-1}$ for $i \in \overline{\mathbb{Z}}_+$, assuming that $t_{i,k}$ was a non-increasing function of i. We assign a non-positive cost of $-\bar{\lambda}$ to each unit of energy left in storage at stage N (i.e. $\mathbf{E}[J_N(s_N)] = -\bar{\lambda} s_N$ for $s_N \leq \bar{s}$), where $\bar{\lambda}$ could, as a reasonable choice, denote the mean of the price distribution (for the remainder of this thesis, it is assumed that $\bar{\lambda}$ denotes the mean of the price distribution). Also, a very small (negative) value is assigned to the thresholds for $i \geq n$ at stage N, to make sure we will not exceed the storage capacity in this stage. Hence, $t_{i,N}$ is a non-increasing function of i, and $t_{i+1,N} \leq t_{i,N}$ is satisfied. It is now time to verify that $t_{i+1,k-1} \leq t_{i,k-1}$ for $i \in \mathbb{Z}_+$ assuming that $t_{i,k}$ was a non-increasing function of i. So, we want to verify:

$$
\begin{aligned}
&t_{i,k}(1 - F_\Lambda(t_{i,k})) + t_{i+2,k}(F_\Lambda(t_{i+2,k})) - h_{i+1,k} \\
&\qquad + \sum_{t_{i+2,k} < \theta \leq t_{i,k}} \theta P_\Lambda(\theta) \;\leq\; t_{i-1,k}(1 - F_\Lambda(t_{i-1,k})) + t_{i+1,k}(F_\Lambda(t_{i+1,k})) + \\
&\qquad\qquad \sum_{t_{i+1,k} < \theta \leq t_{i-1,k}} \theta P_\Lambda(\theta) - h_{i,k}
\end{aligned}
$$

Knowing that $h_{i+1,k} \geq h_{i,k}$, we can remove $-h_{i+1,k}$ and $-h_{i,k}$ from both sides. Then, by taking the negative terms from each side to the other side to make all the terms positive on both sides, and writing the cumulative distribution functions as summations, the above can be rewritten as:

$$
\begin{aligned}
&t_{i,k} + t_{i-1,k} \sum_{\theta=\lambda_{min}}^{t_{i-1,k}} P_\Lambda(\theta) + t_{i+2,k} \sum_{\theta=\lambda_{min}}^{t_{i+2,k}} P_\Lambda(\theta) \\
&\qquad + \sum_{t_{i+2,k} < \theta \leq t_{i,k}} \theta P_\Lambda(\theta) \;\leq\; t_{i-1,k} + t_{i,k} \sum_{\theta=\lambda_{min}}^{t_{i,k}} P_\Lambda(\theta) + \\
&\qquad\qquad t_{i+1,k} \sum_{\theta=\lambda_{min}}^{t_{i+1,k}} P_\Lambda(\theta) + \sum_{t_{i+1,k} < \theta \leq t_{i-1,k}} \theta P_\Lambda(\theta)
\end{aligned}
$$

By taking all the summations to the right hand side and taking $t_{i,k}$ to the left hand side of the inequality, breaking each summation into disjoint intervals, and factoring all the terms in the same interval and merging them into one summation, the above can be rewritten as:

$$\begin{aligned}t_{i-1,k} - t_{i,k} \geq & \sum_{\theta=\lambda_{min}}^{t_{i+2,k}} P_\Lambda(\theta)\left(t_{i-1,k} - t_{i,k} - (t_{i+1,k} - t_{i+2,k})\right) + \\ & \sum_{t_{i+2,k}<\theta\leq t_{i+1,k}} P_\Lambda(\theta)\left(t_{i-1,k} - t_{i,k} - (t_{i+1,k} - \theta)\right) + \\ & \sum_{t_{i+1,k}<\theta\leq t_{i,k}} P_\Lambda(\theta)\left(t_{i-1,k} - t_{i,k}\right) + \sum_{t_{i,k}<\theta\leq t_{i-1,k}} P_\Lambda(\theta)\left(t_{i-1,k} - \theta\right)\end{aligned}$$

We can see by inspection that the above inequality is true. We can see that in the equation on the RHS of the inequality shown above, all the terms that have been multiplied by $P_\Lambda(\theta)$ inside the summations are less than or equal to $t_{i-1,k} - t_{i,k}$; we also know that $\sum_{\theta=\lambda_{min}}^{t_{i-1,k}} P_\Lambda(\theta) \leq 1$. So, given that the summations in the RHS do not overlap, we can clearly see that the RHS equation in the inequality shown above is always less than or equal to $t_{i-1,k} - t_{i,k}$.

The last step is to show, by induction, that the value function is a continuous function. We have defined $\mathbf{E}[J_N(s_N)]$ in such a way that it is convex. We also have defined $h_k(s_k)$ to be convex for all k. Looking at the equations of $J_k(s_k)$ in (3.3) and (3.4), given that $\mathbf{E}[J_{k+1}(s_{k+1})]$ was convex, one would observe that the continuity of $J_k(s_k)$ is trivially satisfied, because v_k^* will be a continuous function of s_k. Similarly, considering the equations obtained for v_k^*, the continuity of $\mathbf{E}[J_k(s_k)]$ for all k also becomes evident and can be easily verified by inspection. This completes the proof.

□

Remark 3. *The convex piecewise linear structure of the value function is non-trivial, but compelling. An outcome of this piecewise linear structure is that the corresponding thresholds have a piecewise constant structure, within the same partitions as in the pieces of the value function. The plot shown in Figure 3-1 illustrates an example of how the thresholds vary as a function of state for a given stage. Note the decreasing trend of the thresholds, which was proved above.*

Remark 4. *The upperbound $\overline{s}$ on the storage capacity can be enforced by choosing*

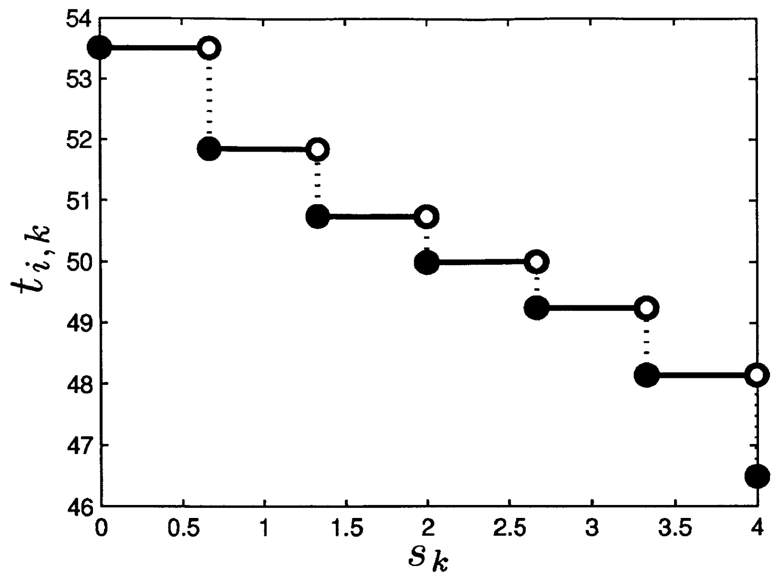

Figure 3-1: An example of how thresholds vary as a function of state for a given stage

$h_{i,k}$ in (3.1) sufficiently large for $i \geq n$, so that it would never be optimal to store energy beyond the capacity $\overline{s}$.

Remark 5. *The $e_{i,k}$ terms in the value function can be derived using the same approach as in the proof of Theorem 1 (described in (3.5) and (3.6)). These terms are given by the following recursive equations, the derivation of which is omitted for brevity:*

$$
\begin{aligned}
&e_{i,N} = 0, \qquad\qquad\qquad i \in 0, 1, 2, ..., n-1 \\
&e_{i,N} = \overline{s}(t_{i,N} - \overline{\lambda}), \qquad i \geq n \\
&\textit{for } k < N \textit{ and } i \in \mathbb{Z}_+ : \\
&e_{0,k} = c_{0,k} + e_{0,k+1}(1 - F_\Lambda(t_{0,k+1})) + \sum_{\lambda_{min} \leq \theta \leq t_{0,k+1}} (e_{1,k+1} + \overline{v}(\theta - t_{1,k+1}))P_\Lambda(\theta) \\
&e_{i,k} = c_{i,k} + e_{i-1,k+1}(1 - F_\Lambda(t_{i-1,k+1})) + \sum_{t_{i,k+1} < \theta \leq t_{i-1,k+1}} e_{i,k+1}P_\Lambda(\theta) + e_{i+1,k+1}F_\Lambda(t_{i,k+1}) + \\
&\qquad \sum_{\lambda_{min} \leq \theta \leq t_{i+1,k+1}} \overline{v}(\theta - t_{i+1,k+1})P_\Lambda(\theta) + \sum_{t_{i+1,k+1} < \theta \leq t_{i,k+1}} (i+1)\overline{v}(\theta - t_{i+1,k+1})P_\Lambda(\theta) \\
&\qquad + \sum_{t_{i,k+1} < \theta \leq t_{i-1,k+1}} i\overline{v}(\theta - t_{i,k+1})P_\Lambda(\theta) + \sum_{t_{i-1,k+1} < \theta \leq \lambda_{max}} \overline{v}(t_{i-1,k+1} - \theta)P_\Lambda(\theta)
\end{aligned}
\tag{3.7}
$$

3.2 Monetary Value of Storage

The baseline for characterization of the value of storage is the cost incurred in the absence of storage, i.e. 0. Therefore, the quantity of interest in this section is how well the consumer does in bringing her expected storage costs below zero. This quantity is the expected monetary value of storage for the consumer, which is denoted by $\mathbf{E}[S]$. Therefore, $\mathbf{E}[S] = 0 - V_0(s_0)$, where $V_k(\cdot)$ is the value function defined in (3.2).

Throughout this section it is assumed that N is fixed, and that $s_0 = 0$, which means that the consumer starts with an empty storage. This implies, using (3.2), that the expected monetary value of storage in this case is:

$$\mathbf{E}[S] = 0 - V_0(0) = -e_{0,0}. \tag{3.8}$$

and its exact value can be found using the recursive equations in (3.7).

For convenience, let us define the following distributions.

Definition. *A 3-point (high, medium, low) symmetric distribution, is a mixture of impulses, where the high and low prices have probability* $1/4$ *each and the medium price* $m = \bar{\lambda}$ *is also the mean, and has a probability of* $1/2$. *The probability mass function (pmf) of the 3-point symmetric distribution is:*

$$P_\Lambda(\theta) = \begin{cases} 1/4 & ; \quad \theta = m - \frac{\Delta}{2} \\ 1/2 & ; \quad \theta = m \\ 1/4 & ; \quad \theta = m + \frac{\Delta}{2} \\ 0 & ; \quad otherwise \end{cases} \tag{3.9}$$

which has a standard deviation of $\sigma = \frac{\Delta}{2\sqrt{2}}$. *Let us also define a discrete uniform distribution that is distributed between non-negative integers a and b, and, for simplicity* *of notation, set $M = b - a + 1$. This distribution has the following pmf:*

$$P_\Lambda(\theta) = \begin{cases} 1/M & ; \quad a \leq \theta \leq b \quad \theta \in \overline{\mathbb{Z}}_+ \\ 0 & ; \quad otherwise \end{cases} \tag{3.10}$$

with standard deviation $\sigma = \sqrt{\frac{M^2-1}{12}}$.

In this section, using the above definition, the following classes of distributions are considered:

- Discrete uniform distribution, with a fixed mean of 50,

- 3-point symmetric distribution, with a fixed mean of 50,

and for each of these distributions, all quantities in the model are fixed, other than σ, the standard deviation of price distribution, and n, the ratio of storage capacity $\overline{s}$ to physical ramp constraint of storage $\overline{v}$. Also, $n \equiv \overline{s}/\overline{v}$ is varied by fixing $\overline{v}$ at $\overline{v} = 10$ and changing $\overline{s}$. Note that as defined in Theorem 1, n only takes on integer values. Using these quantities, let us examine how $\mathbf{E}[S]$ varies as a function of σ and n, once for the case of no storage penalties, and another time in the presence of storage penalties.

3.2.1 Without Storage Penalties

Let us set $h_{i,k} = 0, i \in 0, 1, 2, ..., n-1, \forall k$ so that there is no penalty on storing energy as long as the storage is not filled up. Then, for a fixed time horizon of N=20, let us examine how $\mathbf{E}[S]$ varies with σ and n, for each of the following price distributions:

Discrete uniform prices

Figure 3-2 illustrates how $\mathbf{E}[S]$ changes as a function of σ and n, for a discrete uniform distribution. The plots show that $\mathbf{E}[S]$ increases linearly with σ. As one would expect, $\mathbf{E}[S]$ also increases as the ratio of storage capacity to ramp constraint increases. However, an important observation here is that for a fixed standard deviation, $\mathbf{E}[S]$ becomes almost constant as n increases, starting at a certain value of n. This particular value of n increases as the standard deviation increases. This shows that for a given time horizon, a fixed ramp constraint, and a fixed σ, there exists a certain storage capacity beyond which $\mathbf{E}[S]$ will no longer change noticeably, implying that expansion of the storage capacity beyond that point would practically not bring any further profit to the consumer.

3-point symmetric distribution

A 3-point symmetric distribution was used, and the behavior of $\mathbf{E}[S]$ as a function of σ and n was identical to the case of the discrete uniform distribution. Hence, for conciseness, the corresponding plots are not provided and the reader is referred to the interpretation for the discrete uniform case.

3.2.2 With Storage Penalties

Let us set the storage penalty to $h_{i,k} = 0.1\overline{\lambda}, i \in 0, 1, 2, ..., n-1, \forall k$ as long as the storage is not filled up. Then, for a fixed time horizon of N=20, let us examine how $\mathbf{E}[S]$ varies with σ and n, for each of the following price distributions:

Discrete uniform prices

Figure 3-3 illustrates how $\mathbf{E}[S]$ changes as a function of σ and n, for a discrete uniform distribution. In contrast to the previous case, the plots show that $\mathbf{E}[S]$ is no longer a linear function of σ; however, it is still an increasing function of σ. The compelling observation here is that for a fixed standard deviation, there exists a unique optimal value of n that maximizes $\mathbf{E}[S]$. This optimal value of n increases as the standard deviation increases. This implies that for a given time horizon, a fixed ramp constraint and a fixed σ, there exists an optimal storage capacity. Comparing this result to the case of no penalties with discrete uniform distribution, one would observe that the existence of storage penalties would encourage the consumer to invest in a relatively smaller storage capacity.

3-point symmetric distribution

Figure 3-4 illustrates how $\mathbf{E}[S]$ changes as a function of σ and n, for a 3-point symmetric distribution. The 3-D plot of $\mathbf{E}[S]$ vs. σ and n, and also, the 2-D projection plot portraying $\mathbf{E}[S]$ vs. σ for this case are very similar to the ones for the discrete uniform distribution with storage penalties shown in Figure 3-3. As illustrated in Figure 3-4, the notable difference between the results for this distribution and the discrete uniform case is that the unique optimal value of n that maximizes $\mathbf{E}[S]$ is not necessarily a non-decreasing function of the standard deviation. Comparing this result to the case of no penalties with 3-point symmetric distribution, one would observe that the existence of storage penalties would discourage the consumer from investing in large storage capacity, even if the prices are quite volatile.

3.3 Price-Elasticity

The previous sections characterized a model of an individual consumer's optimal policy for managing storage and the induced monetary value of storage. This section will introduce a simple model of aggregation, in which each individual uses the storage management model. This model is based on giving consumers randomized initial states, and then, computing their consumption, and clustering these as a function of the real-time prices, in order to characterize the *price-elasticity* of demand. In other words, the objective is to study the expected aggregate consumption as a function of price.

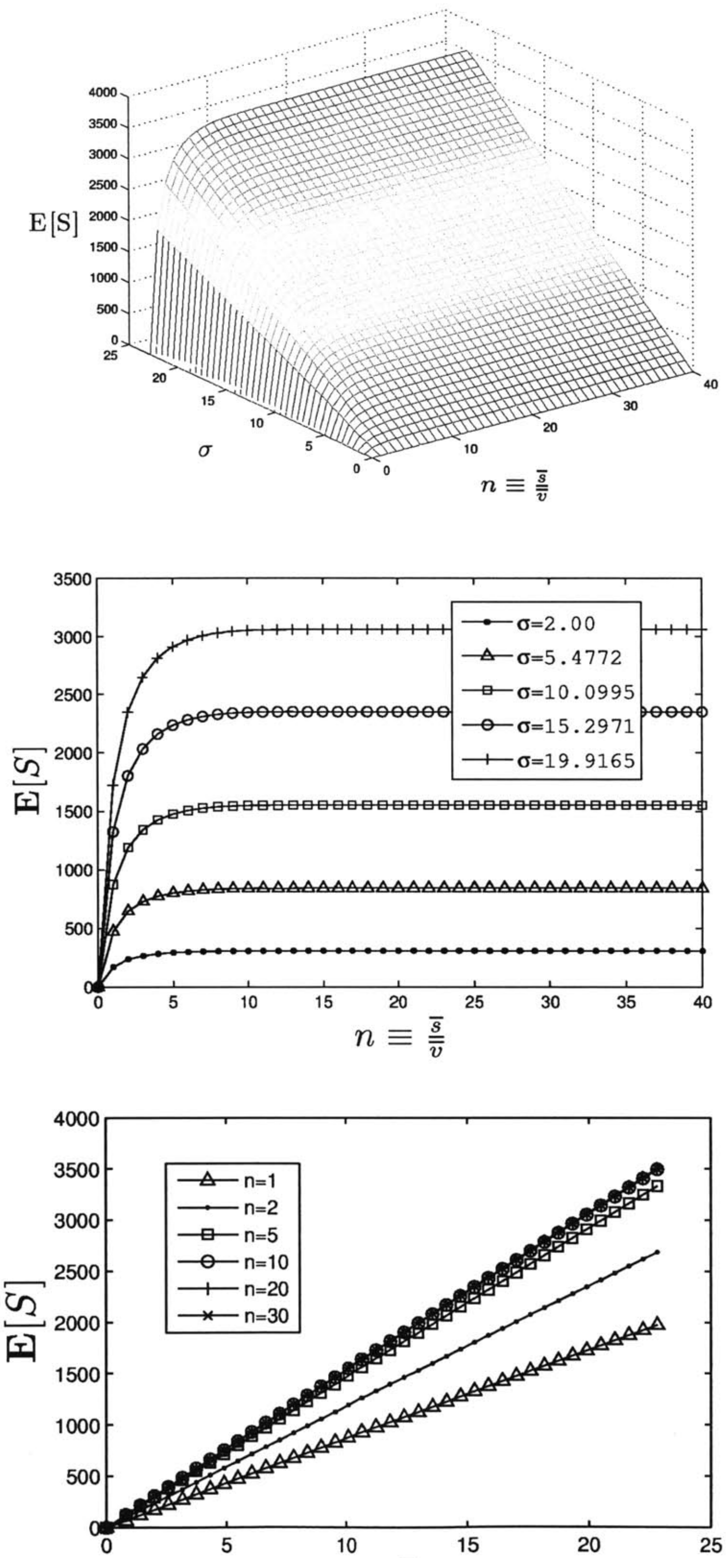

Figure 3-2: The expected monetary value of storage vs. standard deviation of prices and capacity to ramp constraint ratio, in 3-D (top), and the corresponding 2-D projections for a few samples (middle and bottom plots), for the case of no storage penalties, using a discrete uniform distribution.

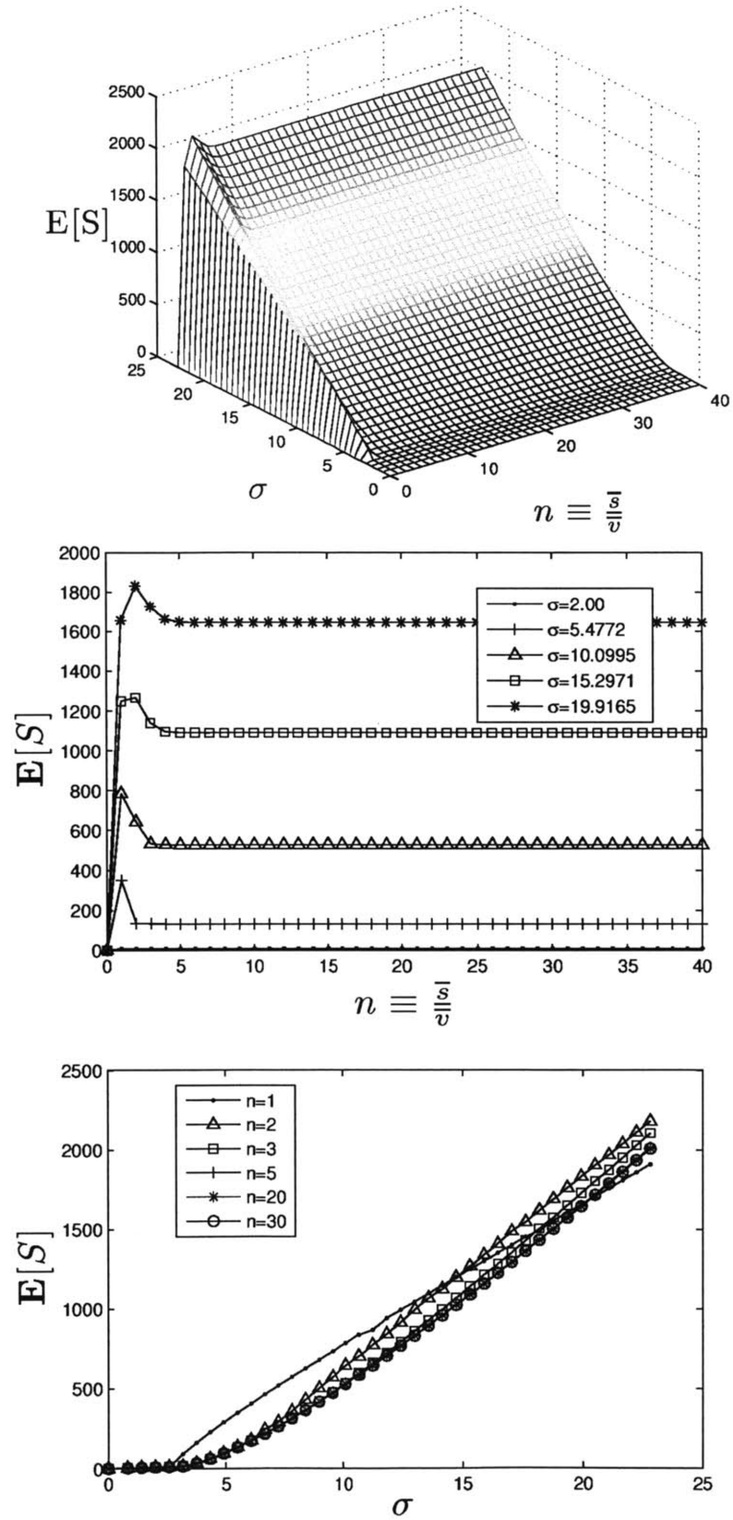

Figure 3-3: The expected monetary value of storage vs. standard deviation of prices and capacity to ramp constraint ratio, in 3-D (top), and the corresponding 2-D projections for a few samples (middle and bottom plots), in the presence of storage penalties, using a discrete uniform distribution.

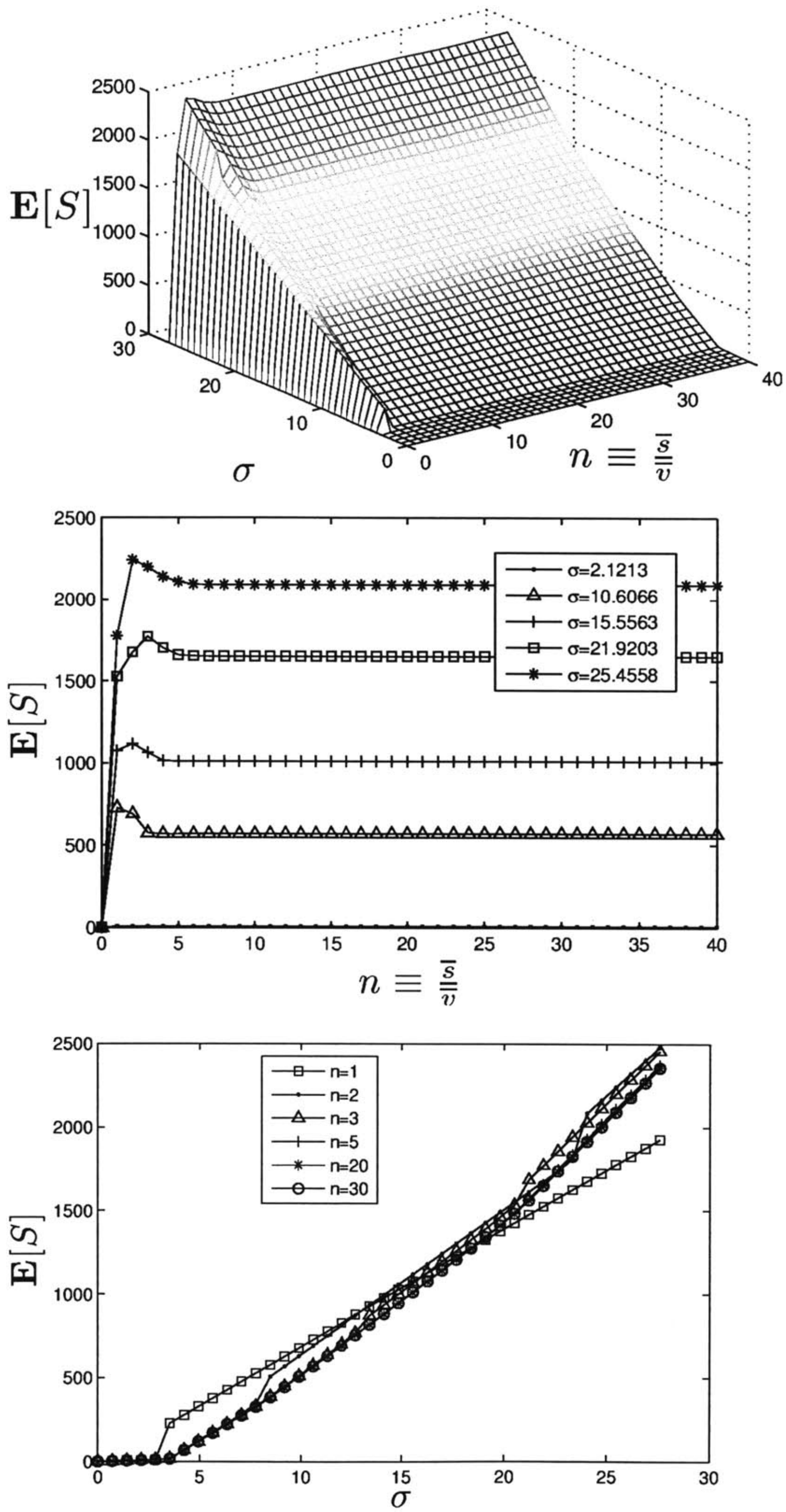

Figure 3-4: The expected monetary value of storage vs. standard deviation of prices and capacity to ramp constraint ratio, in 3-D (top), and the corresponding 2-D projections for a few samples (middle and bottom plots), in the presence of storage penalties, using a 3-point symmetric distribution.

3.3.1 Aggregation Model

The number of consumers is denoted by L, and the aggregation model is specified as follows. Assume that there is a fixed time horizon for all consumers, which is denoted by $\overline{N}$. At every time $k \in \{0, \cdots, \overline{N}\}$, all consumers are given the same price signal λ_k. However, to model random initial states, consumer $j \in \{1, \cdots, L\}$ starts her interaction with the grid at time $k = 0$ with a random amount of energy in her storage that is uniformly distributed over $[0, \overline{s}]$. In other words, at time k, the local state for each consumer is thus s_k^j. The total consumption of all consumers at time k is the ensemble average of the individual v_k^j values.

To make the notion of price-elasticity accurate, one needs a measure of consumption that depends only on price. In the model, however, the consumption depends on price, stage, state, storage capacity, and ramp constraint. Assume the same time horizon for all consumers. In Section 3.2, it was shown that for a fixed time horizon and a fixed standard deviation for a certain distribution, there exists a certain storage capacity beyond which the value of storage will no longer change by much. Also, the physical ramp constraint is usually a known physical parameter. Hence, it is plausible to assume that the storage capacity and physical ramp constraints are fixed for all consumers within the same sector. Moreover, state-dependence can be eliminated by taking expectations. In particular, define:

$$v^j(k, \lambda) = \mathbf{E}_{\lambda_0, \lambda_1, \ldots, \lambda_{k-1}, s_0^j} \left[v_k^{j*} | \lambda_k = \lambda \right],$$

In order to eliminate stage-dependence, think of the consumption-measuring observer as sampling a random time τ uniformly over $\{0, \cdots, \overline{N}\}$. By averaging over this randomness, dependence on prices alone is maintained. More precisely, the quantity of interest is:

$$v_{\text{aggregate}}(\lambda) = \frac{1}{L} \sum_{j=1}^{L} \mathbf{E}_\tau \left[v^j(\tau, \lambda) \right],$$

which is easily captured in numerical simulations by clustering real-time prices, and averaging over each cluster.

3.3.2 Simulations

Aggregation Parameters

In these numerical simulations it is assumed that the number of consumers is $L = 50$, and the average is taken over 20 random instances of prices and consumer initial states. All consumers initiate their interaction with the grid at time $k = 0$ each with a random initial state s_0^j, as described above, and end their interaction with the grid at a global time horizon of $\overline{N} = 288$. This might correspond to, for instance, a period of 24 hours, where real-time prices are updated once in every 5 minutes. They each make optimal utilization of storage according to the proposed model. The physical ramp constraint is set to $\overline{v} = 10$ for all consumers, and a discrete uniform distribution with mean 10 and standard deviation 6.055 is simulated. Using the results in Section 3.2 one would infer that with these model parameters, a storage capacity of $\overline{s} = 50$ is a reasonable choice for all consumers, which is the value used in the simulations for price-elasticity.

In order to investigate how storage penalties affect the price-elasticity of demand, two different scenarios are examined: *(a)* $h_{i,k} = 0 \ \forall k, i < n$ or *(b)* $h_{i,k} = 0.1\overline{\lambda} \ \forall k, i < n$.

Numerical Results

Figure 3-5 illustrates how the aggregate demand changes as a function of price for (a), i.e. when $h_{i,k} = 0 \ \forall k, i < n$

Figure 3-6 illustrates how the aggregate demand changes as a function of price for (b), i.e. $h_{i,k} = 0.1\overline{\lambda} \ \forall k, i < n$.

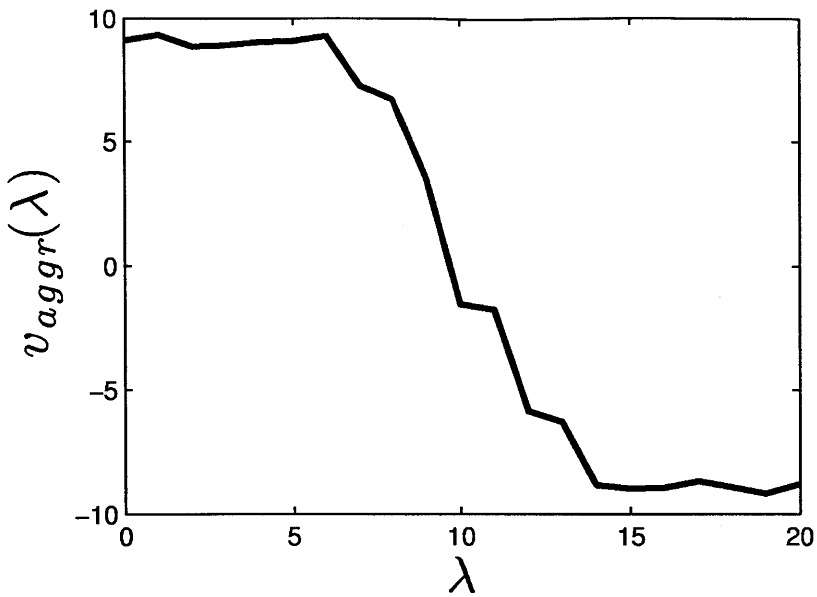

Figure 3-5: Aggregate demand vs. price, with no storage penalties

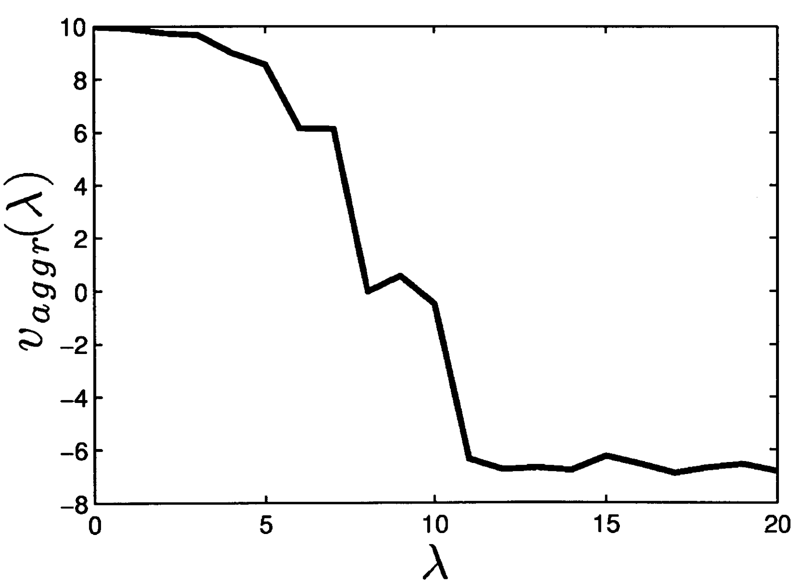

Figure 3-6: Aggregate demand vs. price, with storage penalty.

Interpretation

As the plots for both cases suggest, the aggregate demand for electricity seems to be responsive to prices that fall in the mid-portion of the aggregate demand curve, i.e. those values of λ that are within one standard deviation of the mean. For the case of no storage penalties, this portion serves as a relatively steep transition region, in which the consumer quickly switches from the "buy it all" policy to the "sell it all" policy. The situation is slightly different when storage penalties are imposed. Considering the cost of storing energy, the "buy it all" region has become narrower, and the transition region has instead become wider. The selling policy has practically been reduced to "sell it all if prices are above average" with the exception that due to the high risk of keeping a lot of energy in inventory, the expected sell-back to the grid has been reduced to almost half of the ramp constraint, for practically any price that is greater than average.

To characterize price elasticity of demand in a more quantitative way, one needs to bear in mind that the overall price elasticity should have the firm component of demand in it. That is, the overall price elasticity depends on the amount of storage relative to the firm demand. More storage yields higher elasticity. Similarly, a smaller firm demand results in higher price-elasticity. For instance, in the above simulations for scenario *(a)*, by setting the firm demand, denoted by d^f, equal to 3 times the ramp constraint (i.e. $d^f = 3\overline{v} = 30$), the average price elasticity of demand would be -0.35, whereas setting $d^f = 5\overline{v} = 50$ yields an average price-elasticity of -0.20. Also, in the above simulations for scenario *(b)*, by setting $d^f = 3\overline{v} = 30$, the average price elasticity of demand would be -0.27, whereas setting $d^f = 5\overline{v} = 50$ yields an average price-elasticity of -0.15. Both of these scenarios support the idea that a lower fixed demand relative to storage level yields higher elasticity. However, one would also observe that the average price-elasticity of demand in this case is lower in the presence of storage penalties.

Although a consumer-aggregate model was presented, because the stochastic behavior of each user is the same, the ensemble average provided in this section is equivalent to a single-user expectation. Moreover, if the storage capacity is relatively small compared to the time horizon, then the initial state s_0 only affects the optimal policy of a consumer for the first few stages. Hence, within a short period after the consumers start their interaction with the grid, the states for all the consumers will become the same.

Chapter 4

The Load-Shifting Problem

4.1 The Optimal Policy

This chapter studies the load-shifting problem defined in (2.12). The optimal policy and the associated value function are analyzed in the first section, and then, in the next two sections, the expected consumer savings and the induced price-elasticity of demand are characterized within the same mathematical framework.

4.1.1 Perfect Information about the Price Distribution

In this section, the consumer's optimal policy is characterized for scheduling the shiftable loads based on principles of dynamic programming. This allows for developing a mathematical model of the intertemporal utility of consumption, induced by this optimal load-shifting. It is assumed in particular that the price distribution is perfectly known.

Definition. *Given a distribution function P_Λ with support over a discrete set $\Theta \in [\lambda_{\min}, \lambda_{\max}] \subset \overline{\mathbb{R}}_+$, the modulated expectation function associated with P_Λ is a concave* *function* $\Gamma_\Lambda : [\lambda_{\min}, \infty) \to \overline{\mathbb{R}}_-$ *defined according to*

$$\Gamma_\Lambda(x) = \sum_{\theta=\lambda_{\min}}^{\min(\lambda_{\max},x)} (\theta - x)\, P_\Lambda(\theta) \tag{4.1}$$

This definition extends naturally to continuous distributions, by replacing the summation with an integral.

Remark 6. *The modulated expectation function can be simplified using the method of summation by parts. For* $x \geq \lambda_{max}$*, we can easily see that* $\Gamma(x)=\bar{\lambda} - x$*. However, for the case of* $\lambda_{min} \leq x \leq \lambda_{max}$*, we need to use summation by parts to simplify* $\Gamma(x)$*. We know that random variable* Λ *is distributed between* λ_{min} *and* λ_{max}*. Let us sort in ascending order the values that random variable* Λ *can take on, and denote the k-th element of this ordered set by* θ_k *(e.g. using this notation,* λ_{min} *will be denoted by* θ_1*). Also, denote by* θ_i *the value in random variable* Λ *that is equal to* x*. Finally, let us define* $g_k \equiv F_\Lambda(\theta_{k-1})$ *so that* $P_\Lambda(\theta_k) = g_{k+1} - g_k$*. Then, we can rewrite* $\Gamma(x)$ *as:*

$$\Gamma(x) = \sum_{k=1}^{i-1} (\theta_k - \theta_i)(g_{k+1} - g_k)$$

which, using summation by parts, can be written as

$$\Gamma(x) = [(\theta_i - \theta_i)g_i - (\theta_1 - \theta_i)g_1] - \sum_{k=1}^{i-1} g_{k+1}(\theta_{k+1} - \theta_k)$$

But $g_1 = 0$ *and* $g_{k+1} = F_\Lambda(\theta_k)$*. Hence, we conclude that*

$$\Gamma(x) = -\sum_{k=1}^{i-1} (\theta_{k+1} - \theta_k) F_\Lambda(\theta_k)$$

Now, if random variable Λ *consists of* N *equally spaced elements on the interval* $[\lambda_{min}\lambda_{max}]$ *(i.e.* $\theta_{k+1} - \theta_k$ *is the same for all k), then we can further simplify the above:*

$$\Gamma(x) = \frac{\lambda_{min} - \lambda_{max}}{N-1} \sum_{k=1}^{i-1} F_\Lambda(\theta_k)$$

Similarly, for the continuous case, using integration by parts, we obtain:

$$\Gamma(x) = -\int\limits_{\lambda_{min}}^{x} F_{\Lambda}(\theta)d\theta$$

Consider the load-shifting problem (2.12) with linear disutility of delay (i.e. $p_k(x_k) = -p_k x_k$) and an infinite $\overline{u}$. The optimal policy for this problem can be obtained using the same approach that was used for the storage problem in the previous chapter. As shown in [20], for the load-shifting problem,

(i) The value function is affine, and is characterized by:

$$V(x_k) = \mathbf{E}[J_k(x_k)] = -t_k x_k + e_k, \qquad x_k \leq 0, \tag{4.2}$$

(ii) The optimal policy is a threshold policy characterized by

$$u_k^* = \begin{cases} 0 & \text{if} \;\; \lambda_k > t_k \\ d_k^s - x_k & \text{if} \;\; \lambda_k \leq t_k \end{cases} \tag{4.3}$$

where, the thresholds t_k can be computed via the following recursive equations:

$$t_N = \lambda_{\max}, \quad t_k = p_k + t_{k+1} + \Gamma_{\Lambda}\left(t_{k+1}\right) \tag{4.4}$$

where, $\Gamma_{\Lambda}(\cdot)$ is the modulated expectation function (4.1).

(iii) The constant terms e_k in (4.2) can be computed recursively via the following equations:

$$e_N = \bar{\lambda} d_N^s, \quad e_k = e_{k+1} + d_k^s(t_{k+1} + \Gamma_{\Lambda}\left(t_{k+1}\right)) \tag{4.5}$$

4.1.2 Partial Information about the Price Distribution

One might argue that perfect information about the price distribution is not available to the user in practice. Therefore, in this section, this assumption is relaxed, and instead, an approximation to the modulated expectation function Γ_Λ is proposed, which embodies the dependence of the optimal policy on the price distribution. As shown in [20], if we let $\lambda_{\min} = 0$ and $\lambda_{\max} = 1$, then, given a mean $\mu \in [0,1]$ and an achievable variance σ^2, with $\mathcal{P}$ as the set of all distributions supported on $[0,1]$ that have mean μ and variance σ^2, we can bound Γ_Λ as follows:

$$\underline{\Gamma}_\Lambda(x) \leq \Gamma_\Lambda(x) \leq \overline{\Gamma}_\Lambda(x), \quad \forall x \in [0,1],$$

where:

$$\overline{\Gamma}_\Lambda(x) = \begin{cases} 0 & ;\ x \in \left[0, \mu - \frac{\sigma^2}{1-\mu}\right] \\ (1-\mu)(\mu - x) - \sigma^2 & ;\ x \in \left[\mu - \frac{\sigma^2}{1-\mu}, \mu + \frac{\sigma^2}{\mu}\right] \\ \mu - x & ;\ x \in \left[\mu + \frac{\sigma^2}{\mu}, 1\right] \end{cases}$$

$$\underline{\Gamma}_\Lambda(x) = \begin{cases} \frac{-\sigma^2}{\sigma^2+\mu^2} x & ;x \in \left[0, \frac{\mu^2+\sigma^2}{2\mu}\right] \\ \frac{-\sigma^2\sqrt{(\mu-x)^2+\sigma^2}}{\sigma^2+(\mu-x+\sqrt{(\mu-x)^2+\sigma^2})^2} & ;x \in \left[\frac{\mu^2+\sigma^2}{2\mu}, \frac{1-\mu^2-\sigma^2}{2(1-\mu)}\right] \\ \frac{-(1-\mu)^2(x-1)}{(1-\mu)^2+\sigma^2} + \mu - 1 & ;x \in \left[\frac{1-\mu^2-\sigma^2}{2(1-\mu)}, 1\right] \end{cases}$$

It was also pointed out that both of these bounds are tight pointwise, in the sense that for every $x \in [0,1]$ there exists a distribution $\overline{P}_\Lambda \in \mathcal{P}$ under which $\Gamma_\Lambda(x) = \overline{\Gamma}_\Lambda(x)$ and another distribution $\underline{P}_\Lambda \in \mathcal{P}$ under which $\Gamma_\Lambda(x) = \underline{\Gamma}_\Lambda(x)$.

Furthermore, it was shown that the bounds for the general case of $[\lambda_{\min}, \lambda_{\max}]$ can be obtained via the following transformation:

$$\begin{aligned} &\overline{\underline{\Gamma}}_\Lambda(x; \lambda_{\min}, \lambda_{\max}, \mu, \sigma^2) = \\ &(\lambda_{\max} - \lambda_{\min})\overline{\underline{\Gamma}}_\Lambda\left(\tfrac{x-\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}}; 0, 1, \tfrac{\mu-\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}}, \tfrac{\sigma^2}{(\lambda_{\max}-\lambda_{\min})^2}\right). \end{aligned}$$

It is worth noting that the upper bound $\overline{\Gamma}_\Lambda$ is piecewise linear, whereas the lower bound $\overline{\Gamma}_\Lambda$ is piecewise linear except in the middle segment. To illustrate the usefulness of these bounds, an abstract notion of worst-case optimality was given, by adopting a pointwise min-max notion of optimality for an approximation $\hat{\Gamma}_\Lambda(x)$ of $\Gamma_\Lambda(x)$ at a given point $x \in [\lambda_{\min}, \lambda_{\max}]$. More precisely,

$$\hat{\Gamma}_\Lambda(x) = \min_{\ell} \max_{P_\Lambda \in \mathcal{P}} |\Gamma_\Lambda(x) - \ell| \,.$$

Then it follows that

$$\hat{\Gamma}_\Lambda(x) = \frac{\underline{\Gamma}_\Lambda(x) + \overline{\Gamma}_\Lambda(x)}{2}. \tag{4.6}$$

Remark 7. *The conclusion that this thesis draws from the above results is that the optimal cost of the load shifting problem is bounded above and below by the cost that results from adopting $\overline{\Gamma}_\Lambda$ and $\underline{\Gamma}_\Lambda$ respectively, instead of Γ_Λ, for the computation of the policy thresholds as in (4.4). This conclusion relies on the fact that the optimal cost is an affine expression of the thresholds with non-negative weights, and that adopting the upper or lower bounds results in larger or smaller threshold values respectively. To illustrate these approximations, consider the case where $\mathcal{P}$ represents distributions over $[0,1]$, with mean $1/2$ and variance $1/12$. In Figure 4-1, Γ_Λ is plotted for the special case of a uniform distribution with these parameters, given the upper and lower bounds over $\mathcal{P}$, as well as the min-max approximation of equation (4.6). Considering the fact that the upper and lower bounds are not too far apart, one would expect that partial information should not drastically change the overall behavior of the consumer and her consumption. This theme will be revisited in the next two sections, where individual savings and aggregate price elasticity will be studied.*

4.2 Consumer Savings

One would expect the ability to shift loads to help the consumer bring down her costs, and do so even more effectively when the volatility of prices, measured by the variance, is high. The results in this section support this intuition and characterize it analytically.

The baseline should be the cost that the consumer incurs if the ability to shift loads is taken away from her. In this case, the consumer is forced to purchase the exact

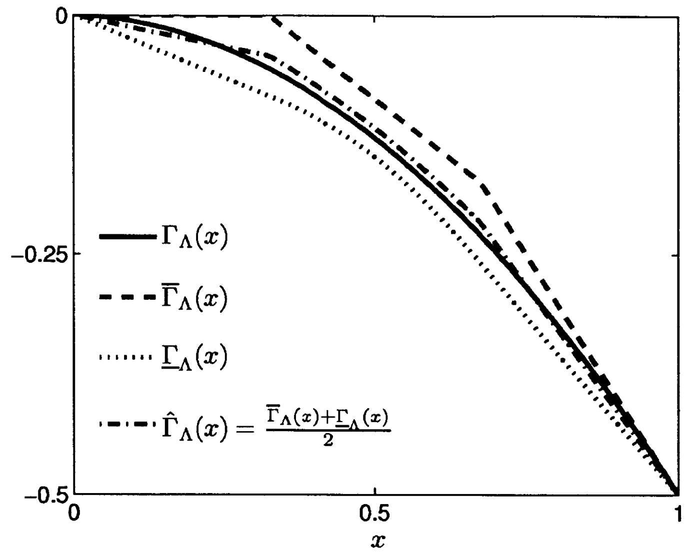

Figure 4-1: Γ_Λ of the uniform distribution, as well as partial information upper and lower bounds, and the min-max approximation, for distributions over $[0, 1]$, with mean $1/2$ and variance $1/12$.

amount of her demand from the grid at each stage. The expected cost of consumption for that consumer will then be $\sum_k d_k^s \bar{\lambda}$. Therefore, the quantity of interest in this section is how well the consumer does in going below this cost. Let us call this the expected savings of the consumer, and denote it by $\mathbf{E}[S]$, which can be defined as follows:

$$\mathbf{E}[S] = \sum_k d_k^s \bar{\lambda} - \mathbf{E}[J_0(x_0)]. \tag{4.7}$$

4.2.1 Savings under Perfect Information

In this section, three known distributions are considered. For simplicity, it is assumed throughout that N is fixed, $x_0 = 0$, $p_k = 0$, and $d_k^s = 1$. Generality is not lost, however, because these terms would principally contribute deterministic scales and shifts, which would not impact the order of dependence that will be derived.

Theorem 2. *Consider the load-shifting problem given in (2.12). Let N be fixed, $x_0 = 0$, $p_k = 0$, and $d_k^s = 1$. Then, for the following classes of distributions*

- *3-point symmetric distributions, fixed mean, and*

- *Discrete uniform distributions, fixed mean,*

all quantities other than the variance σ^2 being constant, we have that $\mathbf{E}[S] = O(\sigma)$.

Proof. Consider the 3-point symmetric distribution. The probability mass function of the 3-point symmetric distribution was defined in (3.9), which has a standard deviation of $\sigma = \frac{\Delta}{2\sqrt{2}}$. In particular, we have $\Delta = O(\sigma)$. Using equations (4.4), (4.2), and (4.5), we have:

$$\mathbf{E}[J_0(x_0)] = e_1 + t_1 + \Gamma_\Lambda(t_1) = e_0$$

Therefore, the total expected cost in this case is equal to e_0. Now, by substituting the recursive equations of e_k (4.5), $1 \leq k \leq N$, into the equation for e_0, we have:

$$\mathbf{E}[J_0(x_0)] = e_0 = t_0 + t_1 + ... + t_{N-1} + \bar{\lambda} \tag{4.8}$$

We know from (4.1) and (4.4) that for $k < N$ and $p_k = 0$,

$$t_k = t_{k+1} + \sum_{\theta=\lambda_{min}}^{\min(\lambda_{\max}, t_{k+1})} (\theta - t_{k+1}) P_\Lambda(\theta) \tag{4.9}$$

Using the distribution of prices as defined in (3.9), and noting that the thresholds are an increasing function of k as defined in (4.9), which implies $t_k < m + \frac{\Delta}{2}$ for $k < N$, we can rewrite (4.9) for $k < N$ as follows:

$$t_k = \begin{cases} \frac{3t_{k+1}}{4} + \frac{m}{4} - \frac{\Delta}{8} & \text{if } \ m - \frac{\Delta}{2} \leq t_{k+1} < m \\ \frac{t_{k+1}}{4} + \frac{3m}{4} - \frac{\Delta}{8} & \text{if } \ m \leq t_{k+1} < m + \frac{\Delta}{2} \end{cases} \tag{4.10}$$

Now note that $t_{N-1} = \bar{\lambda} = m$ is a constant, and therefore

$$t_{N-2} = \frac{m}{4} + \frac{3m}{4} - \frac{\Delta}{8} = m - \frac{\Delta}{8}$$

is $O(-\Delta)$. From (4.10), it follows by induction that t_k is $O(-\Delta)$ for all k. Hence, the total expected cost, shown in (4.8), is also $O(-\Delta)$. It follows that the total expected cost is $O(-\sigma)$, and thus the expected savings are $O(\sigma)$.

Now consider the case of the discrete uniform distribution defined in (3.10). For this distribution, $\sigma = \sqrt{\frac{M^2-1}{12}}$, which means, $M = O(\sigma)$. Also, the prices have mean $\bar{\lambda} = \frac{a+b}{2}$. Hence, $a = \bar{\lambda} - \frac{M-1}{2}$ and $b = \bar{\lambda} + \frac{M-1}{2}$. Using this distribution, we can rewrite (4.9) for $k < N$ as

$$t_k = t_{k+1}(1 - \frac{1}{M}(\lfloor t_{k+1} \rfloor - \bar{\lambda} + \frac{M+1}{2}))+ \frac{1}{M}(\frac{\lfloor t_{k+1} \rfloor(\lfloor t_{k+1} \rfloor + 1)}{2} - \frac{(\bar{\lambda} - \frac{M-1}{2})(\bar{\lambda} - \frac{M+1}{2})}{2}) \quad (4.11)$$

Therefore, as shown in (4.11), considering that $M \geq 1$ and assuming that t_{k+1} was $O(-M)$, we can see by induction that t_k is $O(-M)$. To complete the induction, we need to note that $t_{N-1} = \bar{\lambda}$ is a constant, and

$$t_{N-2} = \bar{\lambda}(1 - \frac{1}{M}(\lfloor \bar{\lambda} \rfloor - \bar{\lambda} + \frac{M+1}{2})) + \frac{1}{M}(\frac{\lfloor \bar{\lambda} \rfloor(\lfloor \bar{\lambda} \rfloor + 1)}{2} - \frac{(\bar{\lambda} - \frac{M-1}{2})(\bar{\lambda} - \frac{M+1}{2})}{2})$$

is $O(-M)$. Hence, the total expected cost shown in (4.8) is also $O(-M)$. Considering that $M = O(\sigma)$, we conclude that the total expected cost is $O(-\sigma)$. Thus, the expected savings are $O(\sigma)$.

□

Remark 8. *For the purpose of comparison, consider a continuous exponential distribution. Such distribution has equal mean and standard deviation, and hence, it is not possible to fix the mean and vary σ. To characterize the case of this distribution, let us define an exponential distribution as follows:*

$$f_\Lambda(\theta) = \begin{cases} \alpha e^{-\alpha\theta} & , \ \theta \geq 0 \\ 0 & otherwise \end{cases} \quad (4.12)$$

which has mean and standard deviation equal to $\frac{1}{\alpha}$ *(where we have* $\alpha > 0$*). We can rewrite (4.9) for the exponential distribution using an integral (for* $k < N$*):*

$$t_k = t_{k+1}(1 - (1 - e^{-\alpha t_{k+1}})) + \int_0^{t_{k+1}} \theta \alpha e^{-\alpha\theta} d\theta$$

which simplifies to

$$t_k = \frac{1}{\alpha}(1 - e^{-\alpha t_{k+1}}) \tag{4.13}$$

Let us look at the last four thresholds for $k < N$*:*

$$\begin{aligned} t_{N-1} &= \bar{\lambda} \\ t_{N-2} &= \tfrac{1-e^{-1}}{\alpha} \\ t_{N-3} &= \tfrac{1-e^{-1+e^{-1}}}{\alpha} \\ t_{N-4} &= \tfrac{1-e^{-1+e^{-1+e^{-1}}}}{\alpha} \end{aligned}$$

As can be seen above, given that t_{N-1} *is a constant, the* $(1 - e^{-\alpha t_{k+1}})$ *term in (4.13) will be only a function of* k *and is independent of* α *for all* k*. Hence, we can rewrite (4.13) as follows:*

$$t_k = \frac{a_k}{\alpha} \quad \textit{for} \quad k \le N-2 \tag{4.14}$$

where $a_k = 1 - e^{-\alpha t_{k+1}}$ *is independent of* α *because the* α *in the exponential term is always cancelled out by the* α *in the denominator of* t_{k+1}*. Therefore, as shown in (4.14), assuming that* t_{k+1} *was* $O(\alpha^{-1})$*, we can see by induction that* t_k *is* $O(\alpha^{-1})$*. To complete the induction, we need to note that* $t_{N-1} = \bar{\lambda}$ *is a constant, and* $t_{N-2} = (1 - e^{-1})\,\alpha^{-1}$ *is* $O(\alpha^{-1})$*. Hence, the total expected cost shown in (4.8) is also* $O(\alpha^{-1})$*. Now, consider the equation for the expected total savings defined in (4.7). Using this equation and the assumptions of this subsection, we have:*

$$\mathbf{E}[S] = \frac{N}{\alpha} - e_0$$

where, using (4.8) and (4.14), we obtain

$$e_0 = \frac{1}{\alpha}(2 + \sum_{k=0}^{N-2} a_k)$$

So,

$$\mathbf{E}[S] = \frac{1}{\alpha}(N - 2 - \sum_{k=0}^{N-2} a_k) \tag{4.15}$$

As can be seen in (4.15) above, for a fixed time horizon, the expected total savings is $O(\alpha^{-1})$*. Now, if we fix* α *and vary* N*, then, considering that for large values of* N *the sum of* a_k*'s will be negligible relative to* N*, we can see that the expected total savings will be* $O(N)$*.*

4.2.2 Savings under Partial Information

Now consider the case when we only have partial information about the price distribution, in that it lies in a set $\mathcal{P}$ of distributions with known support, mean and variance. We know that the upper and lower bounds discussed earlier bound the consumer's expected cost and, consequently, her expected savings.

The behavior of the expected savings can be illustrated numerically using these bounds. Consider the case when $[\lambda_{\min}, \lambda_{\max}] = [0, 75]$ and the mean is $\bar{\lambda} = 50$. In Figure 4-2 the upper and lower bounds on the expected savings are plotted, as the standard deviation σ varies from 0 to 35. Once again, one can observe that the larger the volatility of the prices, the more the consumer is expected to save, even if she has only partial information. This figure also shows the savings under the min-max approximation, which, whilst not being linear, preserves, to some degree, the order of growth that was found under perfect information.

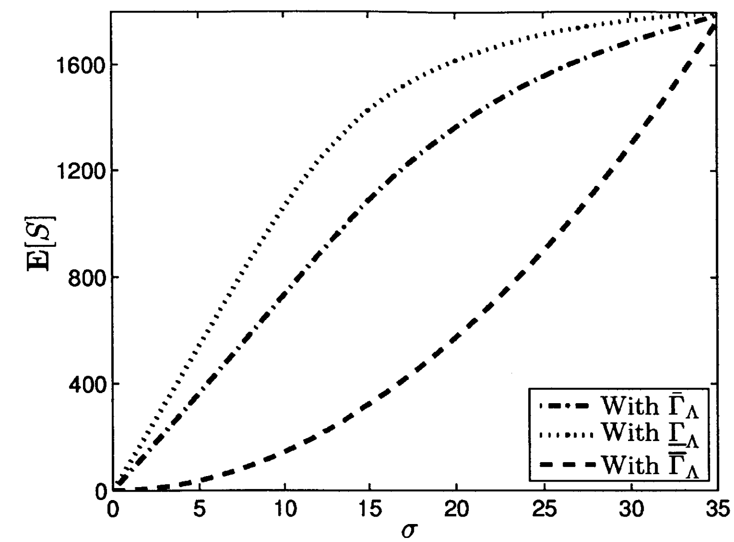

Figure 4-2: Expected savings of a consumer with partial information, as the standard deviation varies.

4.3 Price-Elasticity

In the previous section, an individual consumer's optimal policy for managing shiftable loads was characterized. In this section, a simple model of aggregation will be introduced, where each individual behaves as before. Consumers are given randomized operation periods, and their behavior is simulated. In particular, their consumption is computed over time, and these are clustered as a function of the real-time prices, in order to characterize the *price-elasticity* of demand. In other words, the quantity of interest is the aggregate consumption (u) as a function of price (λ).

4.3.1 Aggregation Model

The number of consumers is denoted by L, and the aggregation model is specified as follows. Assume that there is a global time horizon for all consumers, which is denoted by $\overline{T}$. At every time $t \in \{0, \cdots, \overline{T}\}$, all consumers are given the same price signal λ_t. However, to model random consumption periods, consumer $j \in$

$\{1, \cdots, L\}$ starts consumption at a random time $T_{\text{start}}(j)$ that is uniformly distributed over $\{0, \cdots, \overline{T}-1\}$, and has a random deadline $N(j)$ that is uniformly distributed over $\{1, \cdots, \overline{T}-T_{\text{start}}(j)\}$, conditional on $T_{\text{start}}(j)$. In other words, at time t, the local stage for each consumer is thus $k^j(t) = t - T_{\text{start}}(j)$. However, to simplify notation, indexing is performed by the global time only. The total consumption of all consumers at time t is the ensemble average of the individual u_t^j values, which is taken to be zero if $k^j(t) \notin \{0, \cdots, N(j)\}$.

To make the notion of price-elasticity accurate, one needs a measure of consumption that depends only on price. In the proposed dynamic model, however, the consumption depends on price, stage, and state. State-dependence can be eliminated by taking expectations. In particular, define:

$$u^j(t, \lambda) = \mathbf{E}_{\lambda_0, \lambda_1, \ldots, \lambda_{t-1}, T^j_{\text{start}}, N^j} \left[u_t^{j*} | \lambda_t = \lambda \right],$$

In order to eliminate stage-dependence, think of the consumption-measuring observer as sampling a random time τ uniformly over $\{0, \cdots, \overline{T}\}$. By averaging over this randomness, dependence on price alone can be maintained. More precisely, the quantity of interest is:

$$u_{\text{aggregate}}(\lambda) = \frac{1}{L} \sum_{j=1}^{L} \mathbf{E}_\tau \left[u^j(\tau, \lambda) \right],$$

which is easily captured in numerical simulations by clustering real-time prices, and averaging over each cluster. Although a consumer-aggregate model is presented, because the stochastic behavior of each user is the same, this ensemble average is equivalent to a single-user expectation.

4.3.2 Simulations

Aggregation Parameters

In these numerical simulations it is assumed that the number of consumers is $L = 500$, and the average is taken over 50 random instances of price and consumer arrival. Consumers initiate and end their consumption randomly, as described above, over a global time horizon of $\overline{T} = 720$. This might correspond to, for instance, a period of 24 hours, where real-time prices are updated once in every 2 minutes. To manage their electricity consumption, they each perform optimal load-shifting according to the proposed model.

Load-Shifting Model Parameters

Assume that each consumer starts her consumption with no backlogged demand, i.e. $x_0 = 0$. Also assume that $d_k^s = 1$ for all k.

In order to investigate how load-shifting penalties affect the price-elasticity of demand, two different scenarios are examined: *(a)* $p_k = 0\ \forall k$ or *(b)* $p_k = 0.1\bar{\lambda}\ \forall k$.

Three distributions are simulated: a discretized uniform distribution, the 3-point symmetric distribution defined in (3.9), and a discretized and truncated log-normal distribution, using the same mean $\bar{\lambda} = 0.5$ across distributions.

Numerical Results

Figure 4-3 illustrates how the aggregate demand changes as a function of price for each of the three distributions in scenario (a), i.e. when $p_k = 0\ \forall k$. Each plot contains two graphs; one graph represents the aggregate consumption for the load-shifting problem (2.12) where consumers have perfect information about the price distribution and the other represents the same quantity where users substitute the true threshold function with the min-max optimal approximation (4.6) under partial information.

Figure 4-4 illustrates how the aggregate demand changes as a function of price for each of the three distributions in scenario (b), i.e. when $p_k = 0.1\bar{\lambda}\ \forall k$. As before, each plot contains two graphs, corresponding to perfect or partial information about the price distribution.

Interpretation

In the absence of load-shifting penalties (i.e. when $p_k = 0\ \forall k$), the aggregate consumption varies as a relatively smooth function of price. One can readily interpret this as the aggregate demand for electricity being highly price-elastic.

However, when even a small penalty is assigned to load-shifting, (i.e. when $p_k = 0.1\bar{\lambda}$ $\forall k$), the price-elasticity of demand decreases and the shape of aggregate demand graphs reduce to the policy of purchasing from the grid only when prices are below $\bar{\lambda}$. This means that the price-elasticity of demand is very small nearly everywhere, except when the price is close to a certain threshold, where the demand shows significant elasticity.

The actual amount of price-elasticity would depend on the ratio of shiftable loads to firm loads, which is time-varying and depends on the temporal characteristics of the load. However, it follows that if at any given time a certain portion of the overall load is shiftable, ternary pricing could induce a price elasticity that is proportional to this shiftable portion at the middle and lowest prices, and zero at the highest price. This implies that in situations where shiftable loads comprise a substantial portion of the overall load, a considerable price-elasticity of demand is expected at the middle and lowest prices.

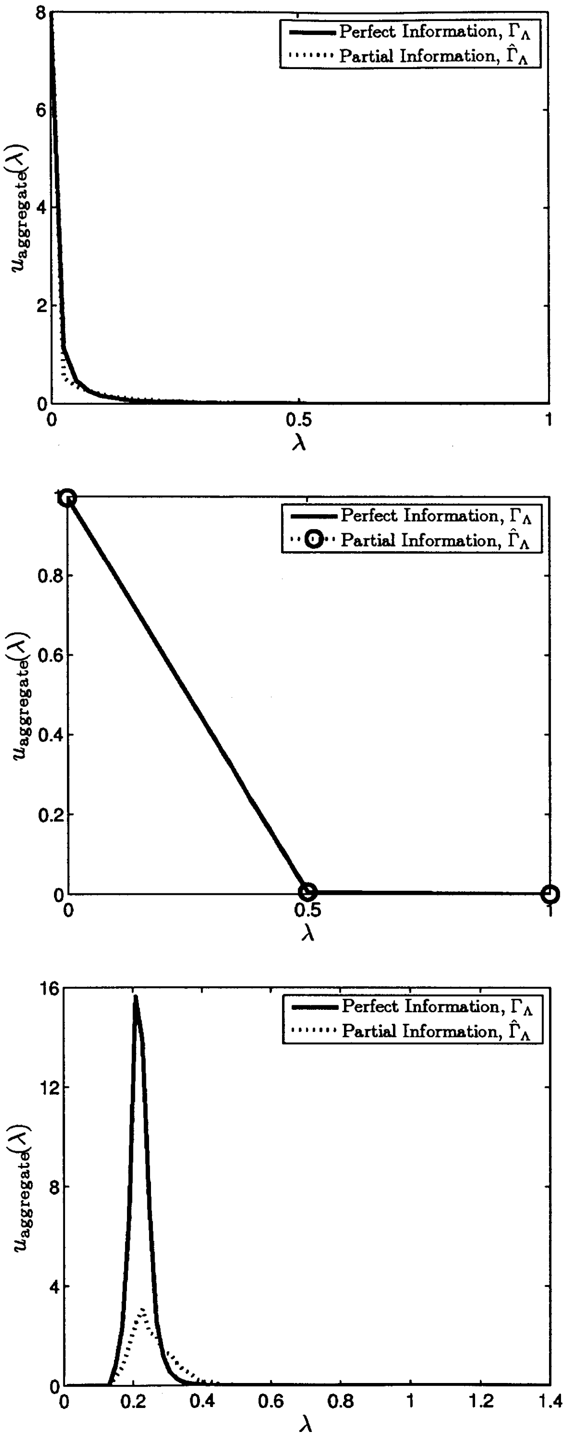

Figure 4-3: Aggregate demand vs. price, with no backlog penalties, for the uniform (top), 3-point symmetric (middle), and truncated log-normal (bottom) distributions.

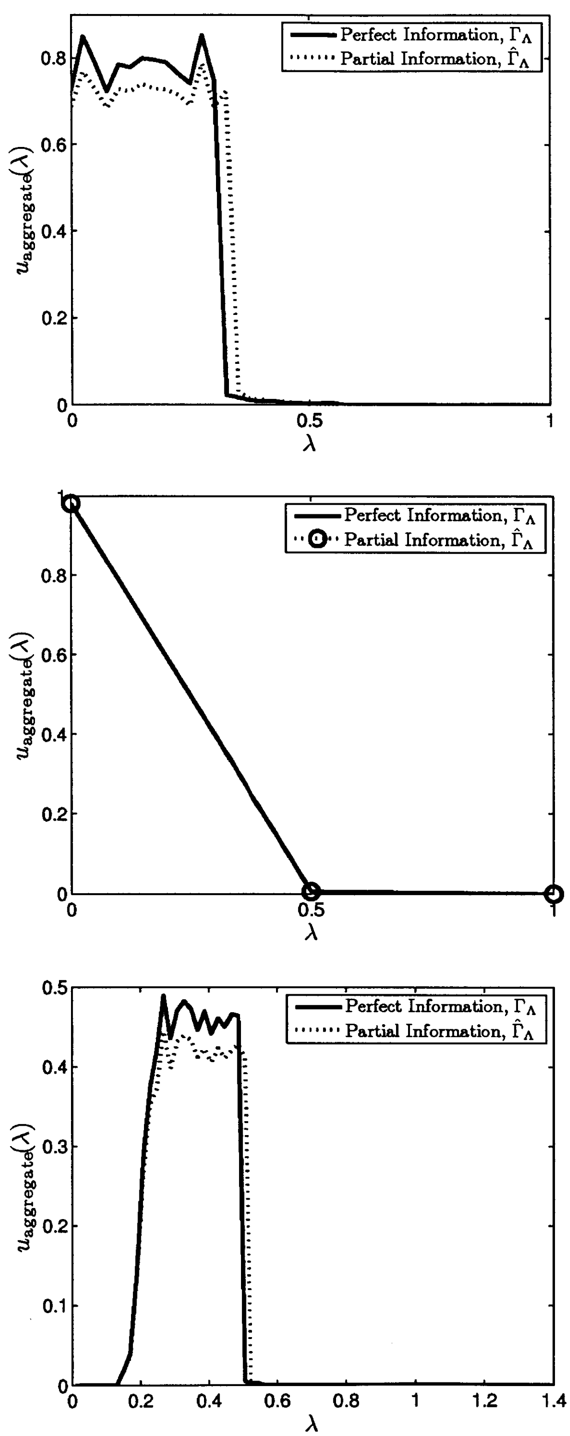

Figure 4-4: Aggregate demand vs. price, with backlog penalty, for the uniform (top), 3-point symmetric (middle), and truncated log-normal (bottom) distributions.

Chapter 5

Conclusions

This thesis presented a dynamic model of intertemporal utility of consumption in response to stochastically-varying electricity prices. It also provided a characterization of price-elasticity of demand created by optimal utilization of storage and the flexibility to shift certain loads to periods of lower prices. The approach was based on analytical characterization of the consumer's optimal policy through a finite horizon stochastic dynamic program. A general model of consumer behavior was first presented, which combined both load-shifting and storage. The model was then decoupled into two subproblems, one for load-shifting and the other one for storage, and each subproblem was studied separately. The proposed models and the findings regarding consumer behavior and price-elasticity can be very useful to various entities, including transmission system operators, utility companies or distribution grid control centers.

On Storage

For the dynamic model of storage management, optimal threshold policies, including analytical expressions for the corresponding thresholds were derived. It was shown, as a very compelling finding, that the value function is a convex piece-wise linear function of the storage state and analytical expressions for this value function were obtained.

Analytical expressions were also provided for the expected monetary value of storage for the consumer. Moreover, it was shown that the expected monetary value of storage increases with the volatility of the prices: larger price variance results in higher expected monetary value of storage. It was also shown that when the ramping rate is finite, the value of storage saturates quickly as the capacity increases, regardless of price volatility.

Finally, it was shown that optimal utilization of storage can induce a considerable amount of price-elasticity of demand. An immediate observation that one could make in these results is that if all the consumers optimally schedule their utilization of storage capacity in the presence of bi-directional meters, a considerable amount of power will be fed back into the grid when the prices are above the mean price. This implies that the consumers' utilization of storage capacity may need to be regulated by the system operator to maintain system balance and stability.

On Load-Shifting

The analytical expressions of the optimal threshold policies and the value function for this problem were used to characterize the total expected savings induced by optimal load-shifting; this characterization was performed using perfectly known price distributions, and also, using analytical bounds that were used to give an approximation to the thresholds in the realistic case when the price distribution is not perfectly known, but rather only its support, mean, and variance are given. It was reported that the total expected savings induced by implementing the optimal load-shifting policies could be considerable, and would increase with volatility of the prices.

Perhaps a deeper finding is that although the demand for electricity is often deemed to be highly inelastic, the introduction of load-shifting mechanisms and the ability to optimally reassign loads to later times can induce a considerable amount of price-elasticity. The characterization of price-elasticity also highlighted the fact that an individual consumers response to a price signal is dependent on both the price and the internal state of the consumer. From the system operation point of view, this dependency on the internal state raises new challenges for adoption of real-time pricing schemes.

Bibliography

[1] K.H. Ahlert and C. van Dinther. "Definition of an optimization model for scheduling electricity storage devices." *Proceedings of Web 2008 - 7th Workshop on e-Business*, December 13, Paris, France, 2008.

[2] ———— and ————. "Sensitivity analysis of the economic benefits from electricity storage at the end consumer level." *Proceedings of the IEEE PowerTech* (28.6.-2.7.), 2009.

[3] C.H. Bannister and R.J. Kaye. "A rapid method for optimisation of linear systems with storage." *Operations Research*, Vol. 39, No. 2, pp 220-232, 1991.

[4] E. Bitar, A. Giani, R. Rajagopal, D. Varagnolo, P. Khargonekar, K. Poolla, P. Varaiya. "Optimal contracts for wind power producers in electricity markets." *Proceedings of 49th IEEE Conference on Decision and Control*, Atlanta, GA, 2010.

[5] J. W. Black, *Integrating Demand into the U.S. Electric Power System: Technical, Economic, and Regulatory Frameworks for Responsive Load*, Ph.D. Thesis, MIT, 2005.

[6] R. N. Boisvert, P. A. Cappers, and B. Neenan, "The benefits of customer participation in wholesale electricity markets." *The Electricity Journal*, vol. 15, no. 3, pp. 41–51, April 2002.

[7] Borenstein, S. "Understanding competitive pricing and market power in wholesale electricity markets." *The Electricity Journal*, vol. 13, no. 6, pp. 49–57, July 2000.

[8] P. D. Brown, J. A. P. Lopes, and M. A. Matos. "Optimization of pumped storage capacity in an isolated power system with large renewable penetration." *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 523531, May 2008.

[9] D. Caves, K. Eakin, and A. Faruqui, "Mitigating price spikes in wholesale markets through market-based pricing in retail markets." *The Electricity Journal*, vol. 13, no. 3, pp. 13–23, Apr. 2000.

[10] P. Constantopoulos, F. C. Schweppe, and R. C. Larson. "A real-time consumer control scheme for space conditioning usage under spot electricity pricing." *Journal of Computers and Operations Research*, vol. 18, no. 8, pp. 751–765, 1991.

[11] A. Faruqui and S. S. George. "The value of dynamic pricing in mass markets." *The Electricity Journal*, vol. 15, no. 6, pp. 45–55, July 2002.

[12] P. Harsha and M. A. Dahleh. "Optimal sizing of energy storage for efficient integration of renewable energy." Submitted to the *50th IEEE Conference on Decision and Control*, Orlando, FL, 2011.

[13] J. H. Kim and W. B. Powell, "Optimal energy commitments with storage and intermittent supply." Submitted to *Operations Research*, 2010.

[14] D. S. Kirschen. "Demand-side view of electricity markets." *IEEE Transactions Power Syst.*, vol. 18, no. 2, pp. 520–527, May 2003.

[15] T. Y. Lee and N. Chen. "Determination of optimal contract capacities and optimal sizes of battery energy storage systems for time-of-use rates industrial customers." *IEEE Transactions on Energy Conversion* 10, 1995, pp. 562-568.

[16] ———— and ————. "Effect of the battery energy storage system on the time-of-use rates industrial customers." *IEEE Proceedings - Generation, Transmission, and Distribution* 141 (5), 1995, pp. 521-528.

[17] D. Livengood and R. Larson, "Energy box: Locally automated optimal control of residential electricity usage." *Service Science*, vol. 1, pp. 1–16, 2009.

[18] D. K. Maly and K. S. Kwan. "Optimal battery energy storage system (bess) charge scheduling with dynamic programming." *IEEE Proceedings -Science, Measurement and Technology*, vol. 142, no. 6, pp. 453-458, November 1995.

[19] A. Papavasiliou and S. S. Oren. "Supplying renewable energy to deferrable loads: algorithms and economic analysis." *IEEE PES General Meeting*, Minneapolis, Minnesota, July 25-29 2010.

[20] M. Roozbehani, A. Faghih, M. I. Ohannessian, M. A. Dahleh. "The intertemporal utility of demand and price-elasticity of consumption in power grids with shiftable loads." Submitted to the *50th IEEE Conference on Decision and Control*, Orlando, FL, 2011.

[21] M. Roozbehani, M. A. Dahleh, S. Mitter. "On the stability of wholesale electricity markets under real- time pricing." *Proceedings of 49th IEEE Conference on Decision and Control*, Atlanta, GA, 2010.

[22] R. Sioshansi, "Increasing the value of wind with energy storage." *The Energy Journal*, vol. 32, no. 2, pp. 130, 2011.

[23] J. M. Yusta and J. A. Dominguez. "Measuring and modeling of industrial demand response to alternative prices of the electricity." *Proceedings of Power Systems Computation Conference*, Session 15, Paper 3, 2002.