Dynamic Pricing of Electric Vehicle Charging Locations: An Application of Optimal Learning

Yu-Sung Huang

Advisor: Professor Warren Powell

April 16th, 2012

Submitted in partial fulfillment
Of the requirements for the degree of
Bachelor of Science in Engineering

Department of Operations Research and Financial Engineering

Princeton University
I hereby declare that I am the sole author of this thesis.

I authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

_______________________________________________
Yu-Sung Huang

I further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research

_______________________________________________
Yu-Sung Huang
To my family,

and all those who believe in me
This thesis, the work over the course of over sixth month, could not have come to fruition without the help of many others.

Within the ORFE department, I would like to thank Professor Powell for his continued guidance during this project. He led me through the framing of this problem and pushed me to think on more theoretical front, which turned out to be an extremely rewarding experience. His advice to repeatedly visualize my data has been critical to my progress and the assurance that my simulations work reasonably. This thesis, by any stretch of imagination, would not have been possible without all of his support, encouragement, and counsel over the last few months.

I would also like to thank Professor Simao for discussing with me the problems of the grid, as well as teaching me how to utilize the computing resources within the PENSA lab, without which my thesis would not have been able to finish on time. Also, thanks to Will Harrell, who volunteered to let me bounce ideas off of regarding the theoretical KG calculations.

In addition, I would like to thank my friends who bore with me these last few months (especially these last few weeks) for providing me constant encouragement. Most of all, I would like to thank my family, who have given me unrelenting support and nudges throughout the course of this project, and have always believed in me regardless of the circumstances.
Abstract

With the introduction of new fuel efficiency requirement for cars, we expect to see a significant increase in the number of plug-in electric vehicles on the road. While such move towards plug-in vehicles reduces U.S. dependency on foreign oil and is expected to be a boon for the environment, such drastic increase in demand for electricity will pose a significant challenge to the already aging and increasingly unreliable U.S. grid. Beyond the concern for generation and transmission, we must also be wary of electric vehicles’ impact on the local distribution grid, especially in large U.S. cities where the distribution grids are often more than half a century old and where car traffic volume will remain high.

In this thesis, we focus on location-of-use demand response policy as a potential solution to this challenge; we conjecture that by using differential pricing for refueling cars at different garages, the local utility can avoid grid failure by redistributing demand. More specifically, we study the early phase of this hypothetical program and focus on how the utility can best learn the demand curve without incurring significant cost. Through multiple simulations, we find that knowledge gradient and interval estimation policies work best in this type of environment.

We end with a discussion of the regulatory context for this type of demand policy program, and bring up key questions that still must be addressed from both a legal and ethical standpoint.
# Table of Contents

Acknowledgement .................................................................................................................. iii

Abstract ...................................................................................................................................... iv

Table of Contents ........................................................................................................................ v

List of Tables .............................................................................................................................. viii

List of Figures ............................................................................................................................. x

Chapter 1:  Introduction.............................................................................................................. 1

1.1.  Wider Context of the Public Policy for Electric Vehicles .............................................. 2

1.2.  Impact on an Aging Grid ................................................................................................. 5

1.2.1.  Electricity Grid Structure and Areas of Concern ....................................................... 6

1.2.2.  Impact of Increasing PHEV Adoption on Electric Grid .............................................. 9

1.3.  Potential Solutions ........................................................................................................... 10

1.3.1.  Supply Side: Upgrading Grid and Storage ................................................................. 11

1.3.2.  Demand Side: Demand Response ............................................................................. 12

1.4.  Overview of the Thesis ..................................................................................................... 13

Chapter 2:  Basic Model ............................................................................................................. 15

2.1.  Basic Outline of the Problem ......................................................................................... 16

2.2.  Basic Parameters and Constants ..................................................................................... 17

2.3.  The Problem in the Context of Optimal Learning ......................................................... 19

2.4.  Modeling the Problem: .................................................................................................... 22

2.4.1.  State Variables ............................................................................................................. 22

2.4.2.  Decision Variables ....................................................................................................... 24
2.4.3. Exogenous Information ................................................................. 24
2.4.4. Transition Functions ................................................................. 25
2.4.5. Objective Function ................................................................. 27

Chapter 3: Learning Policies ................................................................. 29
3.1. Type of Learning Problem and Exploitation vs. Exploration ................. 30
3.2. Policies ......................................................................................... 31
  3.2.1. Pure Exploitation ....................................................................... 32
  3.2.2. Excitation Policy ....................................................................... 33
  3.2.3. Boltzmann Exploration ............................................................... 34
  3.2.4. Interval Estimation ..................................................................... 35
  3.2.5. Knowledge Gradient .................................................................. 36

Chapter 4: Policy Simulation ................................................................. 39
4.1. Garage Locations and Boundaries ....................................................... 39
4.2. Simulation of Demand ...................................................................... 40
  4.2.1. Poisson Arrival Model ............................................................... 41
  4.2.2. Driver Decision-Making ............................................................ 42
4.3. Profit Calculations and Garage Decision-Making ................................. 44
  4.3.1. “Center of Congestion” Approach ............................................. 44
  4.3.2. Garage Price Decision .............................................................. 45
4.4. Generation of Priors and Multiple Simulation Set-Up ............................ 49
4.5. Table of Parameters Chosen ............................................................ 51

Chapter 5: Results .................................................................................... 53
5.1. Simulation Visualization ................................................................. 53
5.2. Importance of Learning .................................................................... 55
5.3. Peculiarities of Bayesian Learning with Parameterized Models ............... 59

5.3.1. Calculation of Correlated Belief Knowledge Gradient - Theory............... 62

5.3.2. Calculation of Correlated Belief Knowledge Gradient - Actual .................. 67

5.3.3. Look-up Table vs. Parametric Beliefs: Some Intuitions......................... 72

5.3.4. Finding "Fulcrum Points" ............................................................................. 76

Chapter 6: Policy Implications ........................................................................... 79

6.1. Impact of Demand Response ......................................................................... 80

6.2. Customer Participation .................................................................................. 82

6.3. Regulatory Framework .................................................................................. 84

6.3.1. Regulation in the Electric Market................................................................. 85

6.3.2. Price-Setting Process for Local Utilities ..................................................... 88

6.3.3. Current Dynamic Pricing Programs.............................................................. 89

6.4. Challenges of Implementing Location-of-Use Demand Response for Electric Vehicles ........................................................................................................ 91

Chapter 7: Conclusions ....................................................................................... 93

References ........................................................................................................... 95
List of Tables

Table 1-1 Survey of Major Fuel Efficiency Policies ................................................................. 4
Table 2-1 Basic Parameters and Constants ............................................................................. 18
Table 2-2 Belief State Demand Parameters .......................................................................... 21
Table 2-3 State Variables ....................................................................................................... 23
Table 2-4 Decision Variables ................................................................................................. 24
Table 2-5 Exogenous Information ......................................................................................... 25
Table 2-6 Transition Functions ............................................................................................. 26
Table 2-7 Contribution Function ............................................................................................ 28
Table 3-1 Pure Exploration ..................................................................................................... 32
Table 3-2 Boltzmann Exploration .......................................................................................... 34
Table 3-3 Interval Estimation ................................................................................................. 36
Table 3-4 Knowledge Gradient .............................................................................................. 38
Table 4-1 Driver Decision Making Process .......................................................................... 42
Table 4-2 Cost of Walking Simulation Variables .................................................................. 43
Table 4-3 Actual Revenue Calculation ................................................................................. 45
Table 4-4 Garage Decision Making ...................................................................................... 47
Table 4-5 Prior Covariance Generation ................................................................................ 50
Table 4-6 Values of Simulation Parameters .......................................................................... 52
Table 5-1 Conversion to Look-Up Table Form ..................................................................... 63
Table 5-2 Calculation of Change in Variance ..................................................................... 64
Table 5-3 KG Function h(a,b) .............................................................................................. 66
Table 6-1 Major Players in Wholesale vs. Retail Markets

................................................................. 86
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Rapid Increase Required in Fuel Efficiency</td>
<td>5</td>
</tr>
<tr>
<td>1-2</td>
<td>Illustration of Electric Industry Structure</td>
<td>7</td>
</tr>
<tr>
<td>1-3</td>
<td>Decreasing Capacity Margin Expected</td>
<td>9</td>
</tr>
<tr>
<td>2-1</td>
<td>Diagram of Demand Response Problem</td>
<td>17</td>
</tr>
<tr>
<td>2-2</td>
<td>Notation Choice for Modeling Time</td>
<td>18</td>
</tr>
<tr>
<td>4-1</td>
<td>Rate of Arrival over the Day</td>
<td>41</td>
</tr>
<tr>
<td>4-2</td>
<td>Myopic Optimization Arrival Pattern</td>
<td>48</td>
</tr>
<tr>
<td>5-1</td>
<td>State of Queens Garages Over Time on Last Day of Summer Season</td>
<td>54</td>
</tr>
<tr>
<td>5-2</td>
<td>Average Cumulative Profit Across Learning Policies</td>
<td>55</td>
</tr>
<tr>
<td>5-3</td>
<td>Average Daily Profit under Different Learning Policies</td>
<td>56</td>
</tr>
<tr>
<td>5-4</td>
<td>Prior Belief and Final Belief Regarding Revenue Curve</td>
<td>58</td>
</tr>
<tr>
<td>5-5</td>
<td>Prior Belief</td>
<td>60</td>
</tr>
<tr>
<td>5-6</td>
<td>Peculiar KG Value (ln)</td>
<td>61</td>
</tr>
<tr>
<td>5-7</td>
<td>Peculiar KG Values</td>
<td>61</td>
</tr>
<tr>
<td>5-8</td>
<td>Theoretical KG Line Plots</td>
<td>65</td>
</tr>
<tr>
<td>5-9</td>
<td>Line Plot for Measurement p = $8.00</td>
<td>67</td>
</tr>
<tr>
<td>5-10</td>
<td>Updated Belief Given Measurement p = $8.00</td>
<td>68</td>
</tr>
<tr>
<td>5-11</td>
<td>Line Plots for Measurement p = $7.75 and $8.25</td>
<td>69</td>
</tr>
<tr>
<td>5-12</td>
<td>Updated Beliefs Given Measurement p = $8.00 and p = $7.75</td>
<td>70</td>
</tr>
<tr>
<td>5-13</td>
<td>Line Plot and Updated Beliefs Given Measurement p = $10.00</td>
<td>71</td>
</tr>
</tbody>
</table>
Figure 5-14 Line Plot for Measurement $p = 10.00$ ................................................................. 72

Figure 5-15 Updating Behavior of Parametric (Linear) vs. Look-Up Table .................... 73

Figure 5-16 Updating Behavior of Parametric (Quadratic) vs. Look-Up Table .......... 75

Figure 5-17 “Fulcrum Point” Terminology Explanation .............................................. 76

Figure 6-1 Probability of Overcapacity Across Different Policies ............................... 81
“Is it a fact -- or have I dreamt it -- that, by means of electricity, the world of matter has become a great nerve, vibrating thousands of miles in a breathless point of time?”

— Nathaniel Hawthorne, The House of the Seven Gables, 1851
Chapter 1: Introduction

The U.S. Federal Energy Regulation Commission estimates that blackouts and disturbances to the electric grid account for up to $180 billion of economic lost every year within the United States.\(^1\) As public policy and increasing environmental consciousness drive the uptake of electric vehicles, the already aging electric grid within the U.S. will bear significant stress. Without steps to ameliorate the impact of the increasing number of electric vehicles, the electric grid will become less and less reliable, which coupled with the necessity of electricity to economic activity, will incur increasingly crippling cost to the U.S. economy.

The below sections first details the wider context that motivated adoption of electric vehicles; secondly, explores impact of electric vehicles on the grid; thirdly, surveys potential solutions, and finally ends with a brief outline of the thesis and its place in this wider context of the challenge on the electric grid.

1.1.  Wider Context of the Public Policy for Electric Vehicles

Crude oil prices have reached unprecedented level over the last decade, rocketing past the $100 per barrel mark at least twice between 2008 and 2012, far surpassing the “historic high” of $36.69 per barrel reached in 1981. Driven by growing demand from developing economies as well as increasing uncertainty of politics in oil-rich Middle Eastern nations, the soaring oil price takes a heavy toll on economies of developed nations, such as the United States. Indeed, expert studies have predicted “a 10% increase in oil prices will result . . . in a level of GDP that is 1.4% lower than it otherwise would be.”\(^2\) Given the current U.S. GDP of roughly $15 trillion, a 10% increase in oil price is then expected to cost the U.S. $210 billion a year. And steady increase in price is not the only concern; increasing volatility of oil prices further act as accomplice to slow economic growth; the Energy Information Agency estimates that about 0.7% of growth was lost between 1997 and 2001 due to the increasing volatility in oil prices.\(^3\) With the political unrest in the Middle East from the Arab Spring movement and an increasingly recalcitrant Iran, oil prices can only be expected to increase, both in terms of absolute price and volatility.

\(^2\) James D. Hamilton, "What is an Oil Shock?" *Journal of Econometrics* 113, no. 2 (2003), 369.
This challenge has not escape the notice of politicians, interest groups, and consumers; as a response to the increasing energy prices, they have helped ushered in few major policy milestones that seek to decrease U.S. dependence on foreign oil, including decreasing gasoline use by transportation sector, seeing as transportation is the end-use sector that consumes the most crude oil, accounting for over 40% of the crude oil consumption in the U.S.\(^4\) Of particular significance are policies that requires automobile makers to increase the fuel efficiency of the cars they produce; these are measured by Corporate Average Fuel Efficiency (CAFE) standards, which is a sales-weighted average of the fuel efficiencies (miles per gallon) of the models a specific company produces. Table 1-1 lists major initiatives regarding increasing fuel efficiency that the U.S. has taken in the past decade:

### Table 1-1 Survey of Major Fuel Efficiency Policies

<table>
<thead>
<tr>
<th>Policy</th>
<th>Impact on Transportation Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bush’s “Twenty by Ten” Initiative</strong></td>
<td>Aims to “to reduce consumption of gasoline by 20 percent over 10 years.”⁵</td>
</tr>
<tr>
<td><strong>Energy Security and Independence Act of 2007</strong></td>
<td>Requires the Corporate Average Fuel Efficiency (CAFE) for the total production models of automobile companies to be at least 35 miles per gallon by 2020.⁶</td>
</tr>
<tr>
<td><strong>Obama’s Executive Fuel Efficiency Agreement in 2011</strong></td>
<td>An agreement with major automobile companies requiring CAFE standard for the whole fleet to be 54.5mpg by 2025.⁷</td>
</tr>
</tbody>
</table>

These increases in the CAFE standard come after a historic lull in regulation, where the standard has not increased for over two decades. The challenge for automobile makers can be visualized below in Figure 1-1. Indeed, for more than two decades, the actual fuel efficiency of the cars sold in the U.S. has hovered right below thirty; hence, drastic increases in fuel efficiency is necessary for automobile companies to reach the 54.5mpg goal by 2025. While the increase in engine performance would certainly help in reaching that goal, automobile company will have to rely heavily on both hybrid electric vehicles and plug-in hybrid electric vehicles (PHEV) to reach this goal, for under the

---


regulation for fuel efficiency, PHEV are rated as much more efficient than a standard vehicle; for instance, current cars average 29.6 mpg\textsuperscript{8}, while PHEV such as Nissan Leaf are rated at an equivalent of 99 mpg.\textsuperscript{9} Hence, the Center for Automotive Research estimated that reaching the goal of 54.5mpg by 2025 would require roughly one in five car sold to be a PHEV (such as Nissan Leaf) in addition to a 35.7% market share by hybrid electric vehicles (such as Prius).\textsuperscript{10}

**Figure 1-1 Rapid Increase Required in Fuel Efficiency\textsuperscript{11}**

![Graph showing CAFE Historic Measures and Future Projection](image)

### 1.2. Impact on an Aging Grid

While increasing adoption of PHEV provides stronger energy independence, not to mention environmental benefits, it is not without ripple effects. Charging a large

---


number of PHEV places significant stress on the electric grid, which may further reduce the reliability of the U.S. grid. The below section first explores basics of how the grid works, and then delves into how increasing adoption of PHEV will pose a significant challenge.

1.2.1. Electricity Grid Structure and Areas of Concern

The electricity industry can roughly be separated along the major steps of generation, transmission, and distribution.12 Figure 1-2 briefly outlines the function of each of these roles within the wider system. Generation entails actually converting other energy sources into electricity, such as coal, gas, nuclear, or wind. Transmission provides conduits that allow electricity to travel over long distances from the plant to the area where electricity is needed. Finally, distribution takes in the power transmitted from the transmission system, brings it down to the correct voltages needed for its customers through substations, and then redistributes to both commercial and residential consumers through a grid that connects each of the customers to the substations.

After the wave of deregulation in electricity industry in the early 1990’s, electricity began to be traded across state borders and in “real-time” markets. While these improvements supposedly increase the economic efficiency of the market, it has also caused electricity prices to become extremely volatile and transactions to become harder to predict. The current reliance on electricity for most of daily economic activities has also made reliability of electricity to be ever more important. And as the Department of Energy states, the current grid system in the U.S., which was mostly established more than four decades ago, is not “designed for the high-level and random nature of electric power transactions that occur today under competitive markets, nor were they designed to

---

provide the high levels of reliability and power quality that many consumers demand today.”

Indeed, on the most macro level, the peak demand for electricity is edging closer to the maximum capacity of the electric system, reducing the size of the margin that is critical to the system’s ability to respond to an increasingly volatile market, which in turn decreases reliability of the system and augment chances of power interruptions. Indeed, based on data from North America Electricity Reliability Company (NERC), Figure 1-3 demonstrates how the margin between capacity and peak demand is expected to decrease significantly in the coming decades. These interruptions and congestions have real impact; indeed, inadequate grid capacity is estimated to cost residents in New York City an extra $90 a year in electricity bills per person, and power interruption, as mentioned at the beginning of the chapter, can cost U.S. up to $180 billion a year. As demand growth continues to outstrip capacity growth, congestion will become more severe, and power interruption will become even more common.

1.2.2. Impact of Increasing PHEV Adoption on Electric Grid

The increasing adoption of PHEV will have drastic impact on the electric grid in terms of significantly augmenting demand. Indeed, when charging, PHEV can consume energy at the rate of a regular household during the peak summer period. Combined with the increase in the number of PHEV over the next decade, PHEV energy usage is expected to increase 1700% by 2020, from 146 thousand MWh to 2.6 million MWh. This increased load is of particular concern, since PHEV tends to be plugged in as people get home and turn on all of their appliances and heating or cooling systems, which will likely amplify the already high peak demand.

While the impact of PHEV on the larger generation and transmission system can be significant, the impact on the local distribution grid is just as concerning. Our current

---

distribution system is not prepared for this level of peak demand; PG&E, the main
electricity distributor on the West Coast, is already expecting this challenge in Berkeley,
California. Indeed, the power grid in most cities of the U.S. have been built gradually
over the last century, and many old equipment are still in use; for instance, at least 5% of
the low-voltage cable in Manhattan was installed before 1930, and ConEd, the utility
company for most of New York City, states that poor management of the local
distribution network can result in “grid failures and possible associated blackouts, fires,
and explosions.”20 In fact, multiple studies have been done to study the reliability of
local distribution grid.21,22 And research at University of Berkeley has recognized that “if
EV loads push peak demand higher, the overall reliability of the grid could be degraded
as a result.”23

1.3. Potential Solutions

Given this significant challenge, the U.S. power industry and government must take
steps to ensure that the electric system will still remain reliable and affordable in the
future. Several potential options are available, which can be roughly categorized into
supply and demand side. Supply side management includes investment in basic
infrastructure and ancillary services, and demand side management is rooted in demand
response programs.

20 Cynthia Rudin et al., "Machine Learning for the New York City Power Grid," IEEE Transactions on
Pattern Analysis and Machine Intelligence 34, no. 2 (2012), 328.
21 Philip Gross, Albert Boulanger, and Marta Arias, "Predicting Electricity Distribution Feeder Failures:
22 Rudin et al., Machine Learning for the New York City Power Grid, 328
23 Nickolas DeForest et al., Impact of Widespread Electric Vehicle Adoption on the Electrical Utility
Business – Threats and Opportunities (2009), 14.
1.3.1. Supply Side: Upgrading Grid and Storage

Upgrading basic infrastructure would be the most straightforward way to meet this challenge. Increasing capacity would allow the margin between capacities and demand to remain at a safe level. However, with the highly deregulated structure today, with many companies competing on different levels, cooperation among these companies to sink significant investment into infrastructure may be difficult. Furthermore, the price tag of such capacity would be even more problematic; the U.S. Energy Information Agency estimates that distribution alone will require $300 billion in investment over the next 15 years. Such price tag often lies outside of the ability of most companies responsible for distribution, since the prices they can charge to customers are often capped. Hence, while increasing capacity is the most straightforward solution, it is neither the most feasible nor the most economically efficient approach.

Another potential solution, storage, takes advantage of the fact that the system approaches capacity only at certain times during the day or year. Indeed, if the supply from low-demand times can be shifted to high-demand times, such as during afternoons or evenings, the impact of peak demand can be much ameliorated. Unfortunately, batteries are quite expensive currently, and hence, as U.S. Department of Energy noted, “electricity cannot be stored economically.”

---

1.3.2. Demand Side: Demand Response

The most promising solution lies in managing demand rather than supply. The core of this approach is to reduce customers’ demand for electricity, and shift demand from peak to non-peak time or location using various incentives. More formally, Federal Energy Regulatory Commission defines demand response as below: 26

*Demand Response is a reduction in the consumption of electric energy by customers from their expected consumption in response to an increase in the price of electric energy or to incentive payments designed to induce lower consumption of electric energy*

By decreasing demand and shifting demand away from peak areas, demand response programs provide a wide array of benefits, according to the U.S. Department of Energy: 27

- **Consumers** – as consumers, reducing one’s own demand and shifting demand to lower-peak times, when the electricity rate may be cheaper, will provide significant savings in electric bills.
- **Market Prices** – if peak demand decreases, then the prices for which the power companies have to bid would be lower, since more expensive generation sources (such as gas-powered engines) will be utilized less often.

---

- **Reliability** – by decreasing peak demand and hence increasing the margin between demand and capacity, demand response gives the system more buffer to deal with volatility in the electricity market, which makes it more reliable, reducing probability of power interruptions.

- **Market Performance** – electricity prices will become less volatile as peak demand decreases, which will promote a more stable market where participants have to pay less to insure against price spikes.

More specifically for the case of PHEV, a potential demand response program would be to employ a location-of-use pricing scheme, where customers would be charged higher prices for charging their vehicle at a location that is experiencing more congestion.

### 1.4. Overview of the Thesis

We chose to use Queens, New York, as a case study to see how potential location-of-use pricing scheme can help redistribute demand geographically to reduce the pressure on the distribution grid in a targeted area; more specifically, the idea is to reduce demand from specific areas experiencing congestion by shifting demand to a less congested area. For instance, it can raise price at a congested area in order to incentivize individuals to charge their cars at other locations. Given the rapidly developing technology, such as smartphone apps, real-time pricing to consumers becomes possible, since they can get advance notice of how much each garage is charging at that specific hour.

Queens is an ideal candidate to study for this issue, for it is in a busy city setting where upgrading grid is less of a feasible option due to high building cost, and there are a good mix of type of demand behaviors across different geographic areas, from places
such as Flushing, which gets heavy day traffic, to more residential areas such as Rockaway, where peak energy usage most likely occur during the night time.

Since such pricing scheme has not been implemented in a city, it is difficult to predict how much price increase utilities should charge to decrease demand without sacrificing profits significantly. This delicate balance between selling electricity while protecting grid reliability under uncertain demand elasticity can be quite a challenge if utilities move towards implementing location-of-use pricing policy for charging electric vehicles. We will tackle this challenge of optimal learning with this thesis.
In order to model the problem, we take the perspective of a power utility (i.e. Consolidated Edison, or ConEd for short) that is responsible for the distribution system in Queens, NY. We make the assumption that parking is concentrated in specific locations within Queens, which we collectively call “garages”; this assumption is relatively reasonable in a busy setting such as Queens, for parking spaces are rare, and it would be much more feasible for utilities to control pricing at a garage level rather than an individual home level. Given this assumption, the challenge becomes how should ConEd price PHEV charging at each garage in order to optimally learn demand and maximize its total end return? The below sections sketch the basic problem, discuss theory of optimal learning, and then further specify the components of the problem by adopting the
framework found in Chapter 5 of *Approximate Dynamic Programming* by Warren Powell, which details the problem with five main components: state variables, decision variables, exogenous variables, transition function, and objective function.\(^\text{28}\)

### 2.1. Basic Outline of the Problem

Figure 2-1 below outlines the basic flow of the specific problem we are modeling. First, the power company projects the load over the grid system in Queens for the whole day excluding the demand from electric vehicles; this is a process that they know how to do reasonably well. Second, given that information, they need to set the prices of charging electric vehicles at each garage on an hour-to-hour basis, one hour ahead of time (hence at 1:00PM, they have to set the prices for cars entering the grid between 2:00PM and 3:00PM). Then, at the end of every hour, they get to observe the number of electric vehicles that have entered the garage within that hour, and the corresponding load on the grid. The behavior of the demand is different for each hour in the day (for instance, rush hour vs. early afternoon), and the goal of the power company is to balance learning how consumer reacts in order to earn as high of a profit as possible without endangering the reliability of the grid. While the choice of time step for the decision-making process is necessarily arbitrary, the hourly time step does capture the reality that the garages cannot shift prices rapidly, and that it must give some advance notice to the consumers of what the prices will be.

---

\(^{28}\) Warren Powell, *Approximate Dynamic Programming*, 2011
For this problem, we chose the length of the learning horizon to be the peak of the summer season – namely July and August – and assume that demand for each hour in the day behaves similarly for each day in the summer season. Hence, the challenge for the utility is to economically learn the demand curve for each hour of the day in order to maximize return over the peak summer season.

### 2.2. Basic Parameters and Constants

Before we dive into modeling, it is useful to specify notational choice as well as certain parameters and constants that will be referred to in the modeling section.

For the purpose of notation, note that in a stochastic dynamic problem involving information, time period $t$ refers to the time period immediately preceding time point $t$, and that any variable with time subscripts less than or equals to $t$ are known quantities at time $t$. Time $t = 0$ corresponds to beginning of the day. This is different than a typical deterministic programming case, where time period $t$ is often considered to be the time period immediately following time point $t$. The time notation used in this thesis can be demonstrated graphically as below in Figure 2-2:
In addition to notation regarding time, we will refer to several parameters and constants listed below in Table 2-1.

**Table 2-1 Basic Parameters and Constants**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>Number of days in the learning horizon, indexed by $d$ $d \in {1, \ldots, D}$</td>
</tr>
<tr>
<td>$T$</td>
<td>Number of hours in each day we are considering, indexed by $t$ $t \in {1, \ldots, T}$</td>
</tr>
<tr>
<td>$N$</td>
<td>The number of garages we have, indexed by $n$ $n \in {1, \ldots, N}$</td>
</tr>
<tr>
<td>$P_{d,t}^n$</td>
<td>Price for a customer to completely charge an EV in garage $n$ who enters during time period $t$ on day $d$. This is charged in addition to the regular parking fee driver pay to the garage owner</td>
</tr>
<tr>
<td>$l$</td>
<td>The average length of time it takes to charge a PHEV; we assume here that all PHEV will remain on the grid for $l$ hours.</td>
</tr>
<tr>
<td>$\alpha_{\text{charge}}$</td>
<td>Is the average amount of power a single PHEV draws from the grid when it is charging</td>
</tr>
</tbody>
</table>
For this problem, it is important to note that while the utility makes decision hour by hour, the time index is done by two subscripts, \( d \) and \( t \). For instance, the 8\(^{th} \) hour during the first day of the summer season would be \( d = 1 \) and \( t = 8 \), and the 18\(^{th} \) hour during the 35\(^{th} \) day of the season would be \( d = 35 \) and \( t = 18 \). Hence, \( t \) subscripts cycles between 0 and \( T \), and \( d \) continue to increase as the season move forward.

### 2.3. The Problem in the Context of Optimal Learning

Optimal learning theory is appropriate for this type of problem, where the decision-maker needs to optimize over a series of decision under uncertainty, and needs to find a cost-and-time-effective way to gain as much useful information as possible. In a basic learning problem, there is a state that we call the “truth,” which is how the demand is simulated, and a state that we call “belief,” which is the best estimate for the true behavior of demand that the utility has, given the information it has received. The goal of the learning problem is to try to learn about the demand effectively so that the “belief” closely matches the “truth.”

Throughout a learning problem, the goal is to gain more accurate estimates of the parameter we desire. There are two main schools within this field – the frequentist view and the Bayesian view. The frequentist school starts with no prior belief regarding the system, and as the experiment goes on, calculates aggregate statistics to estimate the parameters. On the other hand, the Bayesian view starts with a prior distribution and updates it as more data becomes available. We will adopt a Bayesian view, since it is
reasonable to assume that at the summer season, the power company does have some
guess regarding how the demand will behave.

We also know that drivers’ behaviors are different depending on the time of the
day, and we call this information the “state of the world.” Given that, the demand reacts
different to prices under different state of the world, and hence there is a unique “truth”
for each hour, which describes exactly how the demand will respond to prices that the
company sets. Furthermore, this truth is unknown to the company. However, each day,
the company gets to set a price for each of these hours and then observe the result, which
it can then use to continuously “learn” the demand and update its belief regarding the
parameters of the demand behavior.

The demand behavior in this case is not specified by a simple one-dimensional
function, but is a complex mapping of the prices at each garage and the state of the world
to the final number of cars entering each garage; for instance, if a garage charges a higher
price relatively to its neighbors, it is likely to receive less demand, and vice versa.
Ideally, one should be able to model the belief in demand by accounting for all the
possible interaction between demands, which includes any exchange of demand between
garages due to price differentials. However, the current methodologies available in
optimal learning cannot be easily adapted to such complex framework, which would
require modeling a significant number of parameter with interactions that are not easily
accounted for in closed form. Hence, while we use a complex simulation to model the
truth, we model the belief of the demand at each garage as if they are independent of each
other’s pricings. The impact of interactions will be then incorporated into the belief
parameters as each day goes by.
Assuming that it is reasonable to model the demand at each garage as if they are independent of each other, from the optimal learning perspective, we model our belief using a parameterized model, more specifically, quadratic. This assumption is reasonable because the price range of the garage will be limited, and hence while the demand function may take on very different forms outside of the range of prices we consider, the demand function should be relatively well behaved in that narrow price range. Since the final profit is the key objective here, we keep track of our belief using revenue, and then derive a demand curve from that calculation. Hence, we model the belief state regarding the revenue by the following function. Note that we model separate revenue curves for each hour in the day and for each garage (2 dimensions). Table 2-2 further specifies the variables included in this belief state:

\[
\tilde{\alpha}_{d,t}^{n} = \theta_{d,t}^{n,0} + \theta_{d,t}^{n,1} \cdot p_{d,t}^{n} + \theta_{d,t}^{n,2} \cdot (p_{d,t}^{n})^2 + \varepsilon
\]

**Table 2-2 Belief State Demand Parameters**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tilde{\alpha}_{d,t}^{n})</td>
<td>the actual number of vehicles entering garage (n) during hour (t) on day (d) (our observations)</td>
</tr>
<tr>
<td>(\theta_{d,t}^{n,0}, \theta_{d,t}^{n,1}, \theta_{d,t}^{n,2})</td>
<td>constants specifying the demand curve in garage (n) for hour (t) on day (d)</td>
</tr>
<tr>
<td>(p_{d,t}^{n})</td>
<td>is the price of charging for cars entering during hour (t) for garage (n) on day (d)</td>
</tr>
<tr>
<td>(\varepsilon)</td>
<td>is the error term</td>
</tr>
</tbody>
</table>
For numerous numbers of observations, there is a more compact form:

\[ Y_{d,t}^n = (\theta_{d,t}^n)^T X_{d,t}^n + \varepsilon \]

Where:

\[
Y_{d,t}^n = \begin{bmatrix} a_{0,t}^n \\ a_{1,t}^n \\ \vdots \\ a_{d,t}^n \end{bmatrix}, \quad \theta_{d,t}^n = \begin{bmatrix} \theta_{d,t}^{n,0} \\ \theta_{d,t}^{n,1} \\ \theta_{d,t}^{n,2} \\ \vdots \\ \theta_{d,t}^{n,d} \end{bmatrix}, \quad X_{d,t}^n = \begin{bmatrix} x_{0,t}^n \\ x_{1,t}^n \\ \vdots \\ x_{d,t}^n \end{bmatrix}, \quad \begin{bmatrix} 1 & p_{0,t}^n & (p_{0,t}^n)^2 \\ 1 & p_{1,t}^n & (p_{1,t}^n)^2 \\ \vdots & \vdots & \vdots \\ 1 & p_{d,t}^n & (p_{d,t}^n)^2 \end{bmatrix}
\]

Given this belief structure, the challenge of this learning problem is then finding the best \( \theta_{d,t}^n \) for each garage for hour in the day, such that the error between prediction and the observed data is minimal.

2.4. Modeling the Problem:

To further flesh out this model, we adopt the framework found in Chapter 5 of *Approximate Dynamic Programming* by Warren Powell and specify the state variables, decision variables, exogenous information, transition functions, and objective function.

2.4.1. State Variables

State variable is formally defined as “the minimally dimensioned function of history that is necessary and sufficient to compute the decision function, the transition function, and the contribution function.”

\[ \text{Ibid., 159} \]
across time. The state variable of this problem is as below, and Table 2-3 provides further details.

\[ S_{d,t} = (A_{d,t,t'}^n, L_{d,t}, \theta_{d,t}^n, B_{d,t}^n, \Sigma_{d,t}^{\theta,n}) \]

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_{d,t,t'}^n )</td>
<td>keeps track of all the numbers of vehicles that has entered garage ( n ) in the last ( l ) hours ( \forall t \in {0, ..., T} ) ( t' \in {t - l, ..., t} ) ( n \in {1, ..., N} )</td>
</tr>
<tr>
<td>( L_{d,t} )</td>
<td>the state of the Queens grid system over time ( \forall t \in {0, ..., T} ) after taking into account the demand from EV, which allows us to measure the capacity margin at each of the garage at each time</td>
</tr>
<tr>
<td>( \theta_{d,t}^n )</td>
<td>a matrix of the most updated demand parameters for each garage (same ( \theta_{d,t}^n ) in Section 2.3) ( \forall n \in {1, ..., N} ) ( t \in {0, ..., T} )</td>
</tr>
<tr>
<td>( \Sigma_{d,t}^{\theta,n} )</td>
<td>The covariance matrix specifying the covariance between the parameters of ( \theta_{d,t}^n ) ( \forall n \in {1, ..., N} ) ( t \in {0, ..., T} )</td>
</tr>
</tbody>
</table>
2.4.2. Decision Variables

The decision variables of this problem are the prices ConEd should set for each garages for one hour in the future. For instance, at time \( t \), the company has to decide what prices to charge during time period \( t + 2 \). Hence, our decision variable is:

**Table 2-4 Decision Variables**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{d,t+2} )</td>
<td>The price ConEd wishes to charge for each EV that enters garage ( n ) during the time period ( t+2 ) (two hours ahead). ( \forall n \in {1, ..., N} )</td>
</tr>
</tbody>
</table>

2.4.3. Exogenous Information

Two hour after the power company sets the price for each of the charging locations (at time \( t + 2 \)), it gets to observe the number of EVs that have started to charge at each garage within the last hour (time period \( t + 2 \)). This is the only exogenous information that the utility would be able to observe. Table 2-5 below details the list of exogenous information in this problem.
Table 2-5 Exogenous Information

| Notation      | Description                                                                                                                                                                                                 | ∀ | n ∈ {1, ..., N} | ∀ | t ∈ {0, ..., T} |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|----------------|----------------|
| $a_{d,t}^n$   | the number of PHEV that has entered garage $n$ during time period $t$                                                                                                                                            |   |                |                |
| $w_{d,t}$     | the projected weather condition for the day (known at the beginning of each day).                                                                                                                             |   |                |                |
| $L_{d,t}^{grid}$ | the state of the Queens grid system excluding demand from EV for each hour in the day; we assume that the power company is able to predict this at the beginning of every day with perfect precision; hence, for each day at time $t=0$, all the $L_{d,t}^{grid}$ for the rest of the day is a known quantity |   |                |                |

2.4.4. Transition Functions

The updating of most of the state variables is relatively straightforward. The following table specifies updating method for each. For theoretical discussion of the updating method used for variables related to optimal learning under parametric belief $(\theta_t^n, \Sigma_t^{\theta,n})$, please refer to Chapter 8 of *Optimal Learning* by Warren Powell and Ilya Ryzhov (2011). It is important to note that since revenue curves are unique for each hour in the day, the updating time step for our belief state is done day-to-day, while the updating for the grid condition is done hour-to-hour.
Table 2-6 Transition Functions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Updating Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A^n_{d,t+1,t'}$</td>
<td>Simply drop the first row of this matrix and add on $\overline{a^n_{d,t}}$ in the last row.</td>
</tr>
<tr>
<td>$L_{t+1}$</td>
<td>$L_{t+1} = f(L^\text{grid}<em>{d,t+1}, l^n</em>{d,t+1})$</td>
</tr>
<tr>
<td>$\theta^n_{d+1,t}$</td>
<td>$\theta^n_{d+1,t} = \theta^n_{d+1,t} - \frac{1}{y^n_{d+1,t}} \Sigma^n_{d+1,t} \chi^n_{d+1,t} \varepsilon^n_{d+1,t}$</td>
</tr>
<tr>
<td></td>
<td>This updates our predictions about the demand. See below for elaboration.</td>
</tr>
<tr>
<td>$\Sigma^n_{d+1,t}$</td>
<td>$\Sigma^n_{d+1,t} = \Sigma^n_{d+1,t} - \frac{1}{y^n_{d+1,t}} \left( \Sigma^n_{d,t} \chi^n_{d,t} (\chi^n_{d,t})^T \Sigma^n_{d,t} \right)$</td>
</tr>
</tbody>
</table>

More specifically, $l^n_{d,t}$ is calculated as below:

$$l^n_{d,t} = \alpha^{\text{charge}} \sum_{t'=t-l}^{t} a^n_{d,t,t'}$$

In other words, adding up all number of cars still charging in a specific garage, and multiply it by the constant that specifies the amount of power each vehicle draw from the grid when charging. The variables necessary for updating the optimal learning variables are calculated as below.
\[
\begin{align*}
\bullet \ x_{d,t}^n &= [1 \ p_{d,t}^n \ (p_{d,t}^n)^2] & \text{a vector of function of the price} \\
\bullet \ y_{d+1,t}^n &= 1 + (x_{d+1,t}^n)^T \Sigma_{d+1,t}^{\theta,n} x_{d+1,t}^n & \text{a scaling multiplier} \\
\bullet \ \varepsilon_{d+1,t}^n &= \overline{a_{d,t+1}} - \theta_{d,t} x_{d,t}^n & \text{gap between observation and prediction}
\end{align*}
\]

### 2.4.5. Objective Function

The objective of location-of-use demand response is to reduce “hot spots” caused by EV’s in the grid at the location of the garages, which is measured by the amount of margin left at each garage. This is calculated by placing the given energy demand at each garage and running it in a simulation of the grid in Queens in order to produce a mapping of the final load. From the standpoint of the utility, the objective should be to maximize profits, which means that it should sell as much electricity for EV as it can without significant compromise to the grid’s reliability. Roughly speaking, the contribution function should be as below for each hour. Table 2-7 further specifies the significance of individual functions.

\[
C_{d,t}(A_{d,t}^{\text{grid}}, l_{d,t}^{\text{grid}}) = C_{d,t}(l_{d,t}^{\text{EV}}, l_{d,t}^{\text{grid}}) - C_{d,t}(l_{d,t}^{\text{grid}}) + \sum_{n \in \{1, \ldots, N\}} p_{d,t}^n \cdot \overline{a_{d,t}} \\
\forall n \in \{1, \ldots, N\}
\]
Table 2-7 Contribution Function

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{d,t}(A_{d,t}^n, L_{d,t}^{grid})$</td>
<td>The profit or loss the utility experiences during hour $t$ on day $d$</td>
</tr>
<tr>
<td>$C_{d,t}^{EV}(L_{d,t}^{EV}, L_{d,t}^{grid})$</td>
<td>A function that maps the load of the grid to risk in terms of a dollar amount considering the demand from EV’s (if capacity margins are smaller, this cost term is higher)</td>
</tr>
<tr>
<td>$C_{d,t}^{grid}(L_{d,t}^{grid})$</td>
<td>A function that maps the load of the grid to risk in terms of a dollar amount excluding the demand from EV’s</td>
</tr>
<tr>
<td>$\sum_{n\in{1,...,N}} p_{d,t}^n * a_{d,t}^n$</td>
<td>The revenue the utility receives for EV’s that enter the garage during hour $t$.</td>
</tr>
</tbody>
</table>

The overall objective, then, is to maximize the total profit (contribution) the utility company receives during the learning period, which can be summed up as follows.

Section 4.3 provides more details on how we calculated such contributions within our simulation.

$$\max C() = \sum_{d=1}^{D} \sum_{t=0}^{T} C_{d,t}(A_{d,t}^n, L_{d,t}^{grid})$$
Chapter 3: Learning Policies

The classic problem in the field of optimal learning is a somewhat whimsical but notoriously difficult problem of the multi-arm bandit. The scenario can be briefly summarized as follows: the gambler is facing a series of slot machine with different probabilities of winning but with the same exact appearances; by playing a machine and seeing whether he or she wins, the gambler can get a more accurate guess at whether the machine has a high probability of winning or not. In the following sections, we place our problem in parallel with the problem of the multi-arm bandit, and further specify specific policies chosen to deal with this type of problems.
3.1. Type of Learning Problem and Exploitation vs. Exploration

A critical differentiator between learning problems is whether the learning is done offline or online. Offline learning problems are situations where the trials do not incur a cost nor produce profits; on the other hand, online problems are situations where each trial does result in a profit or loss. Often times, offline learning happens in a situation where we are trying to design a system or make a long-term decision, and are given a fix budget to make a specific number of trial runs or experiments. Online learning happens in a situation where we control the system, are able to change decisions day-to-day, and are held accountable to the results of each trial.

In the context of the multi-arm bandit, one can imagine that an offline scenario would be that the gambler is given a certain number of tries to play the machine before he has to select one to play, and that his loss or gain during the trial period would be completely absorbed by the casino. In this case, the gambler should strategize to learn what machine has the highest probability of winning within this number of trials. In the online version, the gambler is playing with his own money, and his goal is to end up with as much winning – or as little loss – as he can; hence, he cannot simply focus on learning which machine has the highest probability of winning; indeed, once he finds a machine that has a pretty good chance of winning, he might not want to test a new machine, since he may suffer a significant loss on the new machine.
Given the setting of our problem, in terms of selling electricity in real-time while still trying to learn the demand curve, our problem can be categorized as an online learning problem.

An online problem has the critical challenge of balancing whether to capitalize on current information we have – which we term exploitation – versus sacrificing the reward for the current period in order to gain new information – which we term exploration. This balance between exploitation and exploration lies at the heart of online-learning problems and accounts for the difficulty of multi-arm bandit problems. Though this problem has not been completely solved, several heuristic policies can be used to provide some guidance into which alternative to try given the amount of information we have.

The goal of this experiment then is to find the policy that is expected to give the utility the best cumulative results at the end of the summer season. Note that the choice of policy only affects what prices we choose next and the amount of information we have regarding the system; it does not change which state variables we need to keep nor the transition functions we use.

### 3.2. Policies

The below section detail policies we have adopted in our simulation. We have chosen not to test the pure exploration policy, which chooses the next price at random. Such a policy does not make much sense in an online learning problem, since it completely ignores the fact that there is cost associated with testing a price that is far from optimal. Also, while it is not included in the below sections, we also used a policy in the control case of not using demand response nor learning, which is to simply charge
a fixed price for all periods – such a policy, which is deterministic, sets the price it is going to test for all iteration before the experiment begins. We also tested a policy that involves no learning – that the belief is not updated from trial to trial. The rest of the policies we consider are sequential policies, in that the information we gained from each trial is then used to inform our decision of what to test for the next trial.

3.2.1. Pure Exploitation

The first policy we test is pure exploitation, which involves choosing the price that gives the maximum return given our current state of the world and our current beliefs. In other words, it chooses the price using the following criteria:

$$P_{d,t+1}^{\text{exp}} = \arg \max_{p \in P} E[\alpha_{d,t+1}^{n} * p - C_{d,t+1}^{n}(A_{d,t+1}^{n}, L_{d,t+1}^{\text{grid}}, p)| P_{d,t+1}^{n} = p]$$

Table 3-1 below further elucidates the terms:

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{d,t+1}^{n}$</td>
<td>The price we are charging for the next period</td>
</tr>
<tr>
<td>$E[\alpha_{d,t+1}^{n}</td>
<td>P_{d,t+1}^{n}]$</td>
</tr>
<tr>
<td>$E\left[C_{d,t+1}^{n}(A_{d,t+1}^{n}, L_{d,t+1}^{\text{grid}}, P_{d,t+1}^{n})\right]$</td>
<td>The expected cost (based on current belief) for charging price $p$ at garage $n$.</td>
</tr>
</tbody>
</table>
It is important to recognize that since we model the belief of each garage separately, these policies also apply on a single-garage basis, in that the choice of price for one garage is completely independent of the choices of price of other garages. For the purpose of future calculations, we also rename the following term, which calculates the expected profit this garage is going to produce given a price, as below:

\[
\mu_{d,t+1}^{n,p} = E\left[ a_{d,t+1}^{n} \ast p_{d,t+1}^{n} \ast -C_{d,t}^{n} (A_{d,t}, V_{d,t}, p_{d,t+1}^{n}) \mid p_{d,t+1}^{n} \right]
\]

Hence, this simplifies the pure exploitation policy as below:

\[
X_{d,t+1}^{exp,n} = \arg \max_{p \in P} (\mu_{d,t+1}^{n,p})
\]

### 3.2.2. Excitation Policy

The pure exploitation policy above does not consider the importance of learning within this problem, in that it may eventually get “stuck” at a sub-optimal price, incurring significant opportunity costs. The excitation policy is a slight modification on the pure exploitation policy that seeks to ameliorate such a problem. The core idea of this policy is that it takes the price chosen by the pure exploitation policy, and then adds or subtracts a random factor from it. This small “perturbation” can reduce the chances of being stuck at a specific place. We simply use a Gaussian Variable with mean 0 and a specific standard deviation:

\[
P_{d,t+1}^{exc,n} = \sigma \ast Z + \arg \max_{p \in P} (\mu_{d,t+1}^{n,p})
\]

It is important to note that excitation is more appropriate for problems with continuous parameters rather than discrete number of choices (such as the multi-arm bandit problem). For in the case of discrete choices, one choice and another, despite their proximity, may have no correlation at all. However, in our case, we can assume that the
demand curve will be continuous over the range of prices, and hence that prices will have significant correlation with each other, and that the closer the prices, the higher the correlation. As such, an excitation policy is appropriate for this problem.

### 3.2.3. Boltzmann Exploration

To be even more sophisticated in balancing the needs of exploration and exploitation, we next try the Boltzmann Exploration method. This method assigns a probability to each choice depending on how good we believe it to be, and then samples randomly from that distribution in order to pick the next price. More specifically, given the $\mu_{d,t+1}$ for each possible price, it assigns probability to each price as follows:

$$P\{p_{d,t+1}^{boltz,n} = p\} = \frac{\exp(\rho \ast \mu_{d,t+1}^{n,p})}{\sum_{p \in P} \exp(\rho \ast \mu_{d,t+1}^{n,p})}$$

Table 3-2 below further specify each term:

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>The Boltzmann constant; we increment this as we move day to day, so that we move from a more exploration approach to a exploitation approach as we get closer to the last day of the learning period.</td>
</tr>
<tr>
<td>$\mu_{d,t+1}^{n,p}$</td>
<td>The expected profit garage $n$ will generate during this time period given price equals to $p$, as discussed in Subsection 3.2.1.</td>
</tr>
</tbody>
</table>
Note that for Botlzmann Exploration, when \( \rho = 0 \), it is the same as pure exploration (since all price gets a equal weight). However when \( \rho \to \infty \), the Boltzmann exploration becomes exactly the same as pure exploitation, for the price with the highest \( \mu_{d,t+1}^{n,p} \) will dominate in the limit. At the beginning of the problem, when there are plenty of chances to capitalize on information learned, it makes more sense to explore, while near the end of the learning season, it makes more sense to do a pure exploitation policy, since we will not have much chance to capitalized on new information. Hence, we increment \( \rho \) as we move from day to day.

### 3.2.4. Interval Estimation

While some of the policy we have explored so far takes into account the balance between exploration and exploitation, none of them have consider the uncertainty for each alternative.

If we are interested in improving the maximum in the long run and are given two prices that have similar \( \mu_{d,t+1}^{n,p} \), it is makes more sense to choose one where our belief has higher uncertainty, for it has a higher chance for improving the maximum than the price for which we have higher level of certainty. In other words, it makes sense to give an “uncertainty bonus” to prices. To be more specific, the interval estimation method is implemented in the following way:

\[
p_{d,t+1}^{i,e,n} = \arg \max_{p \in P} (\mu_{d,t+1}^{n,p} + z_{a} \cdot \sigma_{d,t+1}^{n,p})
\]
where:

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_\alpha$</td>
<td>The interval estimation tuning parameter. This is one of the disadvantages of interval estimation, in that there is a tuning parameter involved; however, past experiment have discovered that a value of 2-3 works reasonably well for most problems.</td>
</tr>
<tr>
<td>$\sigma_{d,t+1}^{n,p}$</td>
<td>The standard deviation of our estimate of $\mu_{d,t+1}^{n,p}$ for price $p$</td>
</tr>
</tbody>
</table>

### 3.2.5. Knowledge Gradient

The most precise of the policy we will utilize for this problem is the knowledge gradient. This policy, just like interval estimation, takes into account the uncertainty of our belief when making choices. The basic aim of this policy is to choose the price that maximizes the expected incremental improvement in our current maximum. In other words, if we define the value of being in the state $S_{d,t}^n$ as $V(S_{d,t}^n) = \max_{p \in P} \mu_{d,t+1}^{n,p}$, then the knowledge gradient chooses the next price as below:

$$p_{d,t+1}^{kg,n} = \arg \max_{p \in P} \left( E \left[ V(S_{d,t+1}^n(p)) - V(S_{d,t}^n) | S_{d,t}^n \right] \right)$$

---

The exact calculation of knowledge gradient value – especially for our case of correlated beliefs under linear regression – can be quite complicated; we leave this discussion to Chapter 5, where we explore the how knowledge gradient value can be effectively zero at certain price points. More importantly, knowledge gradient has three distinct theoretical advantages over other policies we have explored so far:

- No tuning parameters – unlike the interval estimation policy, which requires us to tune $z$, knowledge gradient does not require a tuning parameter.
- Takes into account both uncertainty and correlation – the knowledge gradient policy evaluates each price choice on its potential to improve the overall maximum by forecasting how the correlation between this prices and another will impact our belief regarding the entire demand curve. And similar to interval estimation, it takes into account uncertainty as well.
- Myopically and asymptotically optimal – knowledge gradient is guaranteed to be the optimal policy if we are only going to make one measurement (myopically optimal), but is also guarantee to find the actual optimal price as the number of measurements approach infinity.\(^{31}\)

However, the above is the implementation of knowledge gradient for offline learning, which seeks to maximize the end belief’s maximum, but does not pay attention to the current cost of measuring at price $p$. The online version of knowledge gradient is a simple extension of the offline knowledge gradient (see Powell and Ryzhov’s *Optimal Learning* for derivation); given that there are $D$ total days (and hence $D$ total trials) and that the current day is $d$, the knowledge gradient is computed as below:

\(^{31}\) Ibid, 82.
\[ p_{d,t+1}^{kg,n,online} = \arg \max_{p \in \mathcal{P}} \left( \mu_{d,t+1}^{np} + (D - d)E \left[ V \left( S_{d,t+1}^n(p) \right) - V \left( S_{d,t}^n \right) | S_{d,t}^n \right] \right) \]

While the derivation can run quite technical, this equation has an intuitive interpretation – the Knowledge Gradient optimizes on the next measurement assuming that we only have one more measurement left. The below table explains the meaning of each of the terms:

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_{d,t+1}^{np} )</td>
<td>The expected return for the next period if we choose price ( p )</td>
</tr>
<tr>
<td>((D - d))</td>
<td>The number of measurements we still have to make after the this measurement</td>
</tr>
<tr>
<td>( E \left[ V \left( S_{d,t+1}^n(p) \right) - V \left( S_{d,t}^n \right)</td>
<td>S_{d,t}^n \right] )</td>
</tr>
</tbody>
</table>

Hence, the Knowledge Gradient calculates the return for choosing price \( p \) for the next period, and then adds on the expected improvement from choosing price \( p \) times the number of measurements we will be making. In this way, Knowledge Gradient is acting myopically optimal – that it optimizes the total return of our next measurement.
In this chapter, we lay out the specifics of how we compare the effectiveness of different learning policies. We start off by discussing how we obtained realistic parameters, then detail how we modeled demand, and finally end by sketching how we chose other parameters for our simulation models.

4.1. Garage Locations and Boundaries

In order to provide a more realistic setting for the simulation, we obtained the garage locations by searching on Google Maps and marking down the longitude and latitude of their locations. While an on-the-ground verification and survey of the sizes of the garages would have been ideal, the time limitation of this project makes such
approach infeasible. Hence, we randomly generate the size of each of these garages – in other words, how many electric vehicles each garage can charge – using an exponential distribution with a mean of 40, with any realization less than 20 adjusted to 20. In other words:

\[
size = \max(20, X), \text{ where } X \sim \text{Exp}(50)
\]

From brief research and encounters with garages in Queens, this assumption do not appear to be unreasonable.

In order to simulate the demand within Queens, we also use Google Map to specify the boundary points of Queens, and when we simulate demand, we only consider customers whose final destination is within the proximity of Queens, ignoring the Rockaway area, since there are no garages there, and we assume people who travel there uses other mode of transportation or usually charges at home.

### 4.2. Simulation of Demand

The assumption of location-of-use demand response is that by changing price at each of the garages, the power company can shift demand from one location on the grid to another. We will seek to model the actual behavior of customers with simulations instead of a simple demand curve, since customers in real life switches between garages depending on the price differential, and a simple demand curve that treats each garage as independent of each other would not capture this critical part of how the actual world behaves. We set the number of hours in which drivers can enter garage to be 4:00AM to 11:59PM, assuming that the garages will be shut off between 12:00AM and 4:00AM.
To begin modeling demand, we assume that the arrival of drivers who wish to park in Queens can be modeled reasonably as a Poisson process. Furthermore, one would imagine that the rate of arrival of cars into garages in Queens would differ across time. For instance, arrival rate is likely to be much higher in the morning and the late afternoons, when people are commuting for work, and will most likely be relatively low in the early afternoon, when most people are at work. Hence, we model the arrival process as a non-homogeneous Poisson process, with two peaks in demand – one in the morning and one in the evening. Given that we are accounting for 20 hours in a day, the rate of Poisson process looks similar to this, with a peak around 7:00AM and one around 5:00PM.

Figure 4-1 Rate of Arrival over the Day
### 4.2.2. Driver Decision-Making

Next, to decide whether the drivers would actually drive to Queens, and if so, which garage they would choose, we take the perspective of how a typical driver approaches the choice of parking garages. Given that there is a final destination they wish to reach (their office, a theatre, a restaurant . . . etc.), we assume that the typical driver first find the cost of the garage that is the most ideal for them, and then decide whether to drive into Queens at all. Table 4-1 below specifies this decision-making process in equation form:

**Table 4-1 Driver Decision Making Process**

<table>
<thead>
<tr>
<th>Decision</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g = \arg\min_n [p_{d,t}^n + d^n \ast m^c] )</td>
<td>the customer finds the garage that is cheapest for him after taking into consideration how much cost walking is going to impose on him or her ((d^n)) is the distance from garage (n) to the customer’s final destination, and (m^c) is a multiplier that translates distance into dollar terms; both are customer-specific)</td>
</tr>
<tr>
<td>( I = \begin{cases} 1 &amp; \text{if } p_{d,t}^g + d^g \ast m^c \leq h \ 0 &amp; \text{o.w.} \end{cases} )</td>
<td>the customer then decides whether to take alternative form of transportation ((h)) is the cost of the alternative form of transportation for this customer, which is randomly generated for each customer</td>
</tr>
</tbody>
</table>
The variable $m^c$, which is customer-specific, must take into consideration specific characteristics of the customer or the environment that affect their decision-making. For instance, a wealthy old customer wearing heels while it is raining is much less willing to walk in comparison to a low-income young person in gorgeous spring weather, and hence, $m^c$ in the first scenario would be much higher than in the second scenario. To make the simulation more realistic, we will generate key characteristics of that specific customer, and then calculate the $m^c$ for that customer using the following equation; the specific terms are further specified in Table 4-2.

$$m^c = \overline{m^c} \cdot \left[ 1 + \beta^{inc} Z^{inc} + \beta^{late} Z^{late} + \beta^h Z^h \right] + \beta^{rand} Z^{rand}$$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m^c$</td>
<td>is a constant that is same for all consumers; this is the multiplier for a driver of average income, average lateness, and average health condition</td>
</tr>
<tr>
<td>$Z^{inc}$</td>
<td>is a random variable that represents the income of the driver; it is assumed that if the driver has higher income, then they will be less price sensitive (less willing to walk)</td>
</tr>
<tr>
<td>$Z^{late}$</td>
<td>is a random variable that represents how early or late this driver is relative to their appointment / schedule</td>
</tr>
<tr>
<td>$Z^h$</td>
<td>is a random variable that represents customer’s health condition</td>
</tr>
<tr>
<td>$Z^{rand}$</td>
<td>is a random variable that accounts for all the other factors we have not considered.</td>
</tr>
<tr>
<td>$\beta$</td>
<td>These are scaling constants that are the same for all customers.</td>
</tr>
</tbody>
</table>
4.3. **Profit Calculations and Garage Decision-Making**

While the objective function is specified in Chapter 2, it is not completely clear how to calculate such values, especially the cost term. The revenue term is simply the price charged times the number of arrivals \( p_{d,t}^a a_{d,t} \), but the incremental cost of EV on the grid \( C_{d,t}^{EV} (L_{d,t}^{EV}, L_{d,t}^{grid}) - C_{d,t}^{grid} (L_{d,t}^{grid}) \) is not yet clear.

While ideally we would integrate our model into the realistic Queens grid, the already complicated coding of this model would make run times challenging. Hence, we provide a simplified approach of calculating cost and develop how garages approach picking their next price, namely, how they estimate the amount of profit they will receive if they choose price \( p \) for the next time period.

### 4.3.1. “Center of Congestion” Approach

To prepare for this individual garage approach, we first set a capacity “cap” for each garage, which equals to its total number of EV charging slots times the charge rate \( \alpha^{charge} \). Next, for each day in the simulation, we generate a “center of congestion” within Queens, where we assume that the capacity of the grid is low either due to unexpected high demand from other sources or local grid failure. We then calculate the reduction of capacity at each garage according to its distance from this “center of congestion.” This is done by scaling up a Gaussian probability density function of such a distance, in the form: \( \phi(d) \).
For each garage during each time period, we calculate the profit for garage $n$ during day $d$ and time $t$ by the following equation:

$$C_{d,t}^n(A_{d,t}^n, L_{d,t}^{grid}) = \begin{cases} p_{d,t}^n \cdot a_{d,t}^n - c_{base} \left( D_{d,t}^n \right)^2 & \text{if } D_{d,t}^n < K_d^n \\ p_{d,t}^n \cdot a_{d,t}^n - (c_{base} + c_{surcharge}) \left( D_{d,t}^n \right)^2 & \text{if } D_{d,t}^n > K_d^n \end{cases}$$

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{d,t}^n$</td>
<td>The number of arrivals into garage $n$ in hour $t$ on day $d$</td>
</tr>
<tr>
<td>$K_d^n$</td>
<td>The adjusted “cap” for garage $n$ given its distance to “center of congestion” during day $d$; note that this does not change from hour-to-hour, only from day-to-day</td>
</tr>
<tr>
<td>$D_{d,t}^n$</td>
<td>The power demand from garage $n$ during day $d$ and time $n$. This is: $D_{d,t}^n = \alpha^{charge} \cdot \Sigma_{t'=t-1}^{t} \tilde{a}_{d,t'}^n$</td>
</tr>
<tr>
<td>$c_{base}$</td>
<td>The base cost of electricity the utility pays per kWh</td>
</tr>
<tr>
<td>$c_{surcharge}$</td>
<td>The cost per kWh when over the capacity</td>
</tr>
</tbody>
</table>

### 4.3.2. Garage Price Decision

Given a more concrete contribution function, we can now revisit the variable $\mu_{d,t+1}^{n,p}$ in Chapter 3, which is the profit garage $n$ should expect if it chose price $p$. Indeed, as oppose to a typical learning problem, where $\mu_{d,t+1}^{n,p}$ is based on our beliefs (the expected revenue), the $\mu_{d,t+1}^{n,p}$ in this problem changes day-to-day depending on the
location of the center of congestion, since the cost term changes. We chose calculate
\( \mu_{d,t+1}^{n,p} \) more specifically in the following way:

\[
\mu_{d,t+1}^{n,p} = \begin{cases} 
    p \cdot E[a_{d,t+1}^n] - c^{base}(E[D_{d,t+1}^n])^2 & \text{if } E[D_{d,t+1}^n] < K_d^n \\
    p \cdot E[a_{d,t+1}^n] - (c^{base} + c^{surcharge})(E[D_{d,t+1}^n])^2 & \text{if } E[D_{d,t+1}^n] < K_d^n 
\end{cases}
\]

This is exactly the same as the contribution function above, except for one key
difference, that the demand here is \textit{expected demand}, not actual. A myopic optimization
approach – which tries to maximize profit for the next hour \( t+1 \) – would calculate the
expected arrival \( E[a_{d,t+1}^n] \) and power demand \( E[D_{d,t+1}^n] \) in the following way:

\[
E[a_{d,t+1}^n] = \frac{\left( \theta_{d,t}^{n,w,d0} + \theta_{d,t}^{n,w,d1} \cdot p + \theta_{d,t}^{n,w,d2} \cdot p^2 \right)}{p}
\]

\[
E[D_{d,t+1}^n] = \alpha^{charge} \left[ E[a_{d,t+1}^n] + \sum_{t'=\max(0,t-l+2)}^{t} a_{d,t'}^n \right]
\]
Table 4-4 Garage Decision Making

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha^{charge}$</td>
<td>Rate of charging, namely kWh that a typical EV needs each hour</td>
</tr>
<tr>
<td>$\theta_{d,t}^{n,0}$, $\theta_{d,t}^{n,1}$, $\theta_{d,t}^{n,2}$</td>
<td>Our current belief regarding the revenue for this garage. By calculating the expected revenue and then dividing it by the price chosen, we arrive at the expected number of cars that is suppose to enter the garage during hour $t+1$ given we choose price $p$</td>
</tr>
<tr>
<td>$l$</td>
<td>The typical length of time (in hours) it takes to charge a vehicle</td>
</tr>
<tr>
<td>$a_{d,t}^{n}$</td>
<td>The arrivals for garage $n$ in past hours. This summation term sums up all the cars that will still be in the garage in hour $t+1$ that arrived in hour $t$ or before.</td>
</tr>
</tbody>
</table>

However, this myopic policy does not work well in the larger context for our problem. Any reasonable garage operator knows that during rush hour, it can charge a higher price, and hence it should not lower its price during the hour $t=0$ (4:00AM) to try to fill up his or her garage then. Indeed, if we simply use the myopic optimization policy above, we would have arrival rates like ones below, given that $l = 6$ and the garage can charge 20 cars at a time.
Basically, the garage would fill up at 4:00AM, and these cars will finish charging by 10:00AM, so the garage will not be able to accept any arrivals until then. We can see that such approach to charging garage is definitely sub-optimal, since the garage is not accepting any cars during the peak hours of 7:00AM – 9:00AM and 5:00PM – 8:00PM, when it probably commands a higher price due to increased demand. Such spikes in arrivals are not realistic of how a garage operates.

Hence, we propose to modify our calculations for $E[D_{d,t+1}^n]$ to the following.

$$E[D_{d,t+1}^n] = a^{charge} \left[ \min(1, l - t) E[a_{d,t+1}^n] + \sum_{t' = \max(0, t - l + 2)}^{t} a_{d,t}^n \right]$$

The only major change is that we multiply the expected arrival for the next hour by $\min(1, l - t)$. Hence, for the first 5 hours, the garage is assuming that it already has cars in garages for the last $l - t$ hours. This then help it consider how close to cap it will
be in 6 hours assuming the demand continues to be the same for $t + 1$ through $t + 6$, and hence will not let in as many people as it would otherwise.

We note that this may not be the most optimal policy in choosing prices. However, the focus of this thesis is on optimal learning; a deep exploration of dynamic programming optimization is not within the scope of this thesis; interested reader can find more details regarding dynamic programming optimization problems in *Approximate Dynamic Programming* by Warren Powell.

### 4.4. Generation of Priors and Multiple Simulation Set-Up

The above sections lays out how the “truth” is simulated for this problem. However, the premise of Bayesian learning is also highly dependent on having a “prior belief” from which to update our belief. This prior belief forms the basis for our first measurement decision, and its effect continues throughout the whole simulation, albeit diminishing as we gain more and more information. Hence, how we set our prior belief is critical.

For this problem, we chose a “empirical priors” approach – basically allowing the simulation to test a few points from the truth and form a prior belief from those few observations. However, rather than just letting it test the whole price range, we restricted it to the lower $1/3^{rd}$ of the price range, namely $10.00$ to $20.00$, leaving the range $20.00$ to $40.00$ completely untested. Recalling that the price described here is the price charged in additional to the typical parking lot price, this is quite a reasonable assumption, because before we introduce EV charging, the garages would only have limited experience in the lower price-range (which would be a significant premium on
top of their fee), and would likely not have any experience charging prices that are $20.00 more than their usual market (without EV) price. We utilize this method to provide the $\theta_{0,t}^n$ for each garage and each time period. For the initial covariance matrix, we did not utilize an empirical approach, for there can be cases where charging price between $10.00 and $20.00 will max out the garage, and hence the prior believes that the variance is zero. Hence, instead, we produce the initial covariance matrix of parameters assuming that they are independent with the following equation:

$$
\Sigma_{0,t}^{\theta,n} = \text{diag}[\lambda^W(X^TX)^{-1}]
$$

**Table 4-5 Prior Covariance Generation**

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_{0,t}^{\theta,n}$</td>
<td>The prior belief of the covariance between the parameters $\theta_{0,t}^n$</td>
</tr>
<tr>
<td>$\lambda^W$</td>
<td>The measurement variance</td>
</tr>
<tr>
<td>$X$</td>
<td>$X$ is the design matrix that lists all possible prices, where each row is $[1, p, p^2]$</td>
</tr>
</tbody>
</table>

Also, in order to faithfully test the learning policies, we require multiple trials; for it is possible that one policy would work particularly well in an instance, but not well for most cases. Hence, for each trial, we introduce uncertainty in the following factors:

- Garage size – since we are not able to perform a physical survey of the garage sizes, we use simulation to estimate across multiple possibilities.
- Willingness to walk – the $E[m^c]$ from which all customer $m^c$ are generated from is derived from a normal distribution for each trial.
• Threshold for driving – the parameters from which individual customer
  threshold \( h \) are generated is also similarly derived from a normal distribution
  for each trial.

4.5. Table of Parameters Chosen

The above pages do not detail all the parameters of the simulation. Table 4-6
below summarizes the key parameters; while some of these parameters are taken from
research, most parameters are difficult to find official sources for (such as how much
money people have to save in order to walk an extra block); hence, these parameters are
derived from “reasonable person” standard, in terms of what a reasonable person would
probably behave or do.
Table 4-6 Values of Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D = 60$</td>
<td>We are interested in hours between 4AM and 12AM, and 60 days constitute the critical months of July and August, when electric demand is expected to be the most difficult to manage.</td>
</tr>
<tr>
<td>$T = 20$</td>
<td>$p \in [$10, $40]$ Price charged for EV fueling for a full charge. We discretize this by increments of quarters ($0.25$).</td>
</tr>
<tr>
<td>$E[m^c]$</td>
<td>The average from which all customers’ trade-off ratio between walking and money is generated ($\bar{m}^c$). For each trial, we generate the $E[m^c]$ from a Gaussian(1.5,0.5) distribution, and within each simulation, we generate individual customer’s $\bar{m}^c$ from a Gaussian ($E[m^c],0.5$) distribution.</td>
</tr>
<tr>
<td>$E[h]$</td>
<td>The average from which all customer’s threshold $h$ is derived. For each trial, we generate $E[h]$ from a Gaussian(55,10) distribution, and within each trial, we generate each customer’s threshold $h$ from a Gaussian ($E[h], E[h]/3$).</td>
</tr>
<tr>
<td>$\beta = 0.3$</td>
<td>The coefficient in front of random normal variables for income, lateness, health condition, and a pure random factor for each customer’s $m^c$ calculation. For simplicity, we set these all equals to 0.3.</td>
</tr>
<tr>
<td>$l = 6$</td>
<td>The typical length of time (in hours) it takes to charge a vehicle, and the electric consumption of a charging EV per hour in kWh. $^{32}$</td>
</tr>
</tbody>
</table>

---

Chapter 5: Results

While the parameters of our simulation and our model by no means perfectly match reality, there are qualitative comparisons that can be made to draw some key lessons. In the below section, we start off with some visualization of how the simulation runs, discuss the key results of our simulation, and explain the underlying intuition for our results. In addition, we explore some interesting findings regarding Bayesian learning model with parametric beliefs.

5.1. Simulation Visualization

In order to provide some assurance that the simulation is working properly, we graph the state of Queens garage each hour within each day. Figure 5-1 below
contains graphs demonstrating the inflow of cars into garages in Queens from
4:00AM until 10:00AM for the last day of the summer season for a specific trial.

**Figure 5-1 State of Queens Garages Over Time on Last Day of Summer Season**

Within the blue boundary, which outlines the boundary of Queens, each dot represents the garages we have chosen to model, and they are colored blue if they are under 70% of their capacity, green if within 70% to 130% of capacity, and red if over 130% capacity. We see that the arrival is as one would expect from a typical day: we start off the morning with almost no cars – hence most garages under capacity – and then as we approach rush hour, the garage fills up and gets close to or over capacity. While this does not guarantee that our simulation is completely correct, the above graphic suggests that our simulation produces quite reasonable results.
5.2. Importance of Learning

The purpose of running simulations is to compare the effect of different learning policies, and to see which policy would be the best for a utility approaching this demand response problem without a clear understanding of the demand curve. Figure 5-2 outlines the final results of the simulation, graphing the average cumulative profit the utility would experience over 60 days of the summer season under different policies over 40 trials. The error bars represent 95% confidence interval for each policy (namely, 1.96 times the standard deviation of the mean).

**Figure 5-2 Average Cumulative Profit Across Learning Policies**

![Average Cumulative Profit Across Learning Policies](image)

On the horizontal axis, *ctrl* means the control case policy, where we simply charge the minimum price at all times, *nl* is the no-learning case (a pure exploitation policy without updating of belief), *exp, kg, ie, exc, boltz* means pure exploitation, knowledge gradient, interval estimation, excitation, and Boltmann exploration policies, respectively. Please see Chapter 3 for more detailed explanation on these policies.
We see that knowledge gradient and interval estimation outperformed other policies quite significantly, while Boltzmann exploration performed moderately, consistently breaking even. However, it still performed much better than the control, no-learning, pure exploitation, and excitation policies; all these policy produce negative profits.

To begin to understand the mechanics underlying such result, we examine how much learning each policy has done by tracking the profit across each day. Figure 5-3 below graphs the average daily profit under each policy from the first day \((d = 1)\) of the summer season to the last day of the summer season \((d = D = 60)\).

**Figure 5-3 Average Daily Profit under Different Learning Policies**

<table>
<thead>
<tr>
<th>Average Daily Profits under Different Learning Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>-150</td>
</tr>
<tr>
<td>Average Daily Profit ($1000/day)</td>
</tr>
<tr>
<td>ctrl</td>
</tr>
<tr>
<td>nl</td>
</tr>
<tr>
<td>exp</td>
</tr>
<tr>
<td>kg</td>
</tr>
<tr>
<td>ie</td>
</tr>
<tr>
<td>exc</td>
</tr>
<tr>
<td>boltz</td>
</tr>
</tbody>
</table>

There are two key metrics to use when examining this graph: the “starting” point for the policies, and then how the policy performed better across time. Given that the prior belief does not map the true demand well, these two metrics asks whether the policy
is able to guess a close-to-optimal point right off, and whether it continues to learn for the rest of the season.

In terms of starting points, we see that control policy performs the worst, followed by a tight pack of no-learning, pure exploitation, and excitation policies. Boltzmann exploration and knowledge gradient both manage to start off close to breaking even, while interval estimation at its first measurement usually brings in significant profit. In terms of learning, we see that every policy except for knowledge gradient experience no or little improvement. Knowledge gradient is quickly able to catch up and surpass interval estimation despite starting off at the break-even point.

One might question why interval estimation policy seems to perform worse as the days go one. We hypothesize that this is due to the game-theoretic nature of this optimization problem, which further complicates the learning problem. Recall that the demand curve is not a simple mapping of one garage’s price to its demand, but a complex mapping of all garages’ prices, and hence, a garage’s demand curve can shift depending on the prices of nearby garages. When we perform a learning policy, we assume that all garages individually tries to maximize their cumulative profit at the end. Interpreted in this light, interval estimation is the only policy that has not reached competitive equilibrium. One can almost imagine that in the knowledge gradient scenario, the garages quickly learned that if they all charge a higher price (act as a cartel), they can gain significant profit, while the rest of the policy are “stuck” at a low-price low-margin competitive landscape; interval estimation, then, can be interpreted as garages starting off having great cooperation, but quickly devolving into a price war. While to prove that such behavior accounts for the negative slope of the interval estimation curve is difficult,
we do believe that the above explanation makes some intuitive inroads into explaining this peculiar behavior.

To further examine the reason behind these differences, we also kept track of prices each policy choose to measure as well as the updated demand curves at a specific garage and a specific hour for each trial across the days. To explain how different policies learned differently, we used Figure 5-4, which compares the final belief under each policy, as well as all the prices tested by the policy along with the corresponding observations in revenue. The first frame is the prior belief we started with, which has a different scale:

**Figure 5-4 Prior Belief and Final Belief Regarding Revenue Curve**

We see that out of all the policies, knowledge gradient has a final belief with the highest maximum out of all these curves, while interval estimation performs relatively well. The rest of the policies, however, only tested points close to the minimum, and
hence got “stuck” there, and their belief in the end does not deviate too far from the prior belief.

Note that the above discussion is only for a specific hour for a specific garage under a specific trial, and does not prove that this mechanics accounts for all the difference in the results. However, given the context of the problem of theory of optimal learning, this explanation seems quite plausible, at least at explaining the majority of the differences between the policies.

We have also kept track of probability of over capacity under each policy. This result is presented in Chapter 6, since it relates more closely to public policy implications.

5.3. Peculiarities of Bayesian Learning with Parameterized Models

During the debugging process of our simulation, we found a peculiarity with the behavior of the knowledge gradient (KG) values. Usually, for a model with correlated beliefs, one would expect the KG value graphed against the prices to be relatively smooth and either concave or convex. However, in our case, we found that the KG value exhibits sharp dips at two points, which is quite unexpected considering the fact that we are using a parameterized regression model that enforces a smooth belief curve; it is rather hard to imagine how such smooth curves in demand can result in such drastic changes in KG value. The below figure is the prior belief on which KG is calculated; note that the potential choices of prices range from $2.00 to $10.00, incremented by $0.25.
The KG value graph is as below; note that this is the offline KG model and sets $\mu_{d,t+1}^{n,p}$ as the revenue, without considering the costs. Figure 5-6 graphs the log of the KG values while Figure 5-7 graphs the KG values directly. We see a significant dip at prices $3.00 and $8.00, where the KG value is basically zero – this seems to suggest that there is no benefit to measuring $3.00 or $8.00.
Again this appears to be highly unusual for such smooth graph to produce such
dips in KG value. The next few sections elucidate why KG value behaves this way. We
first start with a general overview of how correlated belief KG values are computed, and
then proceeding to apply the theory to this case study. Finally, we provide some intuition
as to why parametric beliefs behave in this peculiar way.
5.3.1. Calculation of Correlated Belief Knowledge Gradient - Theory

We chose the correlated belief KG with linear regression model in approaching this problem. The linear regression model allows us to keep track of only a few parameters across time rather than large look-up tables of beliefs and covariances. Nevertheless, when calculating the KG value we still utilize the same calculation as that of a look-up table. It is not necessary to get into the detailed mathematics in order to explain the phenomena seen above, and in the discussion below we only lightly touch on mathematics and instead focus more on the intuitions for explanation. Interested reader can see Chapter 5 and Chapter 8 of Powell and Ryzhov’s *Optimal Learning* for more detailed theoretical discussion regarding KG.

In order to calculate the KG value, we must first convert the linear regression model into a look-up table form. The estimated revenue and the complete covariance matrix for each price are developed as below. For the sake of simplicity, we drop all the subscripts except for $d$ in the following discussion, where $d$ represents each day, and hence each trial of updating; these variables are introduced in Chapter 2.

\[
\begin{align*}
E[Y^{d+1}] &= (\theta^d)^TX \\
\Sigma^d &= X(\Sigma^\theta,d)(X)^T
\end{align*}
\]
### Table 5-1 Conversion to Look-Up Table Form

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[Y^{d+1}]$</td>
<td>This is a vector of expected revenue for each given price</td>
</tr>
<tr>
<td>$\theta^d$</td>
<td>A vector of current beliefs regarding the regression parameters</td>
</tr>
<tr>
<td>$X$</td>
<td>The design matrix that outlines each possible choice of price, where each row is composed of $[1\ p\ p^2]$</td>
</tr>
<tr>
<td>$\Sigma^d$</td>
<td>The variance covariance matrix for all the possible choice of $p$. This can be quite a large matrix.</td>
</tr>
<tr>
<td>$\Sigma^{\theta,d}$</td>
<td>The 3-by-3 variance covariance matrix for our parameters $(\theta^d)$. This is what we keep track of instead of the much larger $\Sigma^d$.</td>
</tr>
</tbody>
</table>

In calculating the KG value, we are interested to see if a measurement will change our decision – that is, it would change the maximum choice of our updated belief. To begin considering the impact of one measurement on the overall maximum, we first calculate the *change* in the variance of our belief due to measuring price point $x$. The below equation highlights the calculation of the vector that details the impact of measuring $x$ on our belief for each of the alternatives.

$$
\tilde{\sigma}(x) = \frac{\Sigma^d e_x}{\sqrt{\Sigma^d_{xx} + \lambda W}}
$$
### Table 5-2 Calculation of Change in Variance

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_x$</td>
<td>A column matrix that has all zero entries except for the $x^{th}$ entry. Hence, $\Sigma^d e_x$ gives us the $x^{th}$ column of $\Sigma^d$.</td>
</tr>
<tr>
<td>$\lambda^W$</td>
<td>Our measurement variance. Note that this is the variance of our next measurement $W^d$.</td>
</tr>
</tbody>
</table>

Given that we assume our measurement noise is normally distributed, we can then write our updated belief, given that we chose to get test price $x$, as a vector using the below equation ($Z^{d+1}$ is realization from a random normal variable)

$$E[Y^{d+1}] = E[Y^d] + \sigma(x^d) \cdot Z^{d+1}$$

Now, the key to KG is that we are only concerned with the maximum of our new belief model. Hence, we are only interested in finding what is the expected increase in maximum given our choice to measure $x$. Recognizing that the above equation is actually a system of equations, outlining how our belief regarding each alternative will shift depending on the value of $z$. We can graph each of these equations in a graph similar to the one below (renaming $\theta^d = a$, and $\sigma(x^d) = b$, creating lines in the form $y = a + bx$) and highlight the maximum point at each possible $Z^{d+1}$. The bolded line, which is our top choice at a specific point of $Z^{d+1}$, forms an “upper envelope” to this graph:
Recognize that the upper envelop at $Z = 0$ is the current maximum. Depending on the value of $Z$, several things can happen:

- **if $Z = 0$,** our belief regarding the expected revenue for each price does not change, namely that the vectors $E[Y^{d+1}] = E[Y^d]$, and hence our maximum choice also remains the same: $\arg\max (E[Y^{d+1}]) = \arg\max (E[Y^d])$

- **However, if $Z$ changes but remains between $c_1$ and $c_2$,** our expected revenue changes, which means $E[Y^{d+1}] \neq E[Y^d]$, but our maximum choice remains the same: $\arg\max (E[Y^{d+1}]) = \arg\max (E[Y^d])$

- **In the last case, if the $Z$-value is beyond $c_1$ or $c_2$,** not only will our belief change, $E[Y^{d+1}] \neq E[Y^d]$, but also our maximum choice would change

---

33 Adopted from p.88 of Powell and Ryzhov’s *Optimal Learning*
argmax(E[Y^{d+1}]) \neq argmax(E[Y^d]), deviating from the choice that represented the line \( y_2 = a_2 + b_2 z \) in the above figure.

KG then takes that into account and calculates the expected total increase in our maximum that we can expect from this choice, which is a weighted sum of the values on the upper envelope by the probability of the corresponding \( Z \) value; more specifically, this is calculated by first finding all the break points \( c_i \) as illustrated above and eliminating the lines that regardless of value of \( Z \) cannot become the maximum. Interested reader can read Chapter 5 of Powell and Ryzhov for detailed algorithm of how such breakpoints are calculated. Then, the KG value for choice \( x \) is calculated in the following manner:

\[
h(a, b) = \sum_{i=1}^{M} (b_{i+1} - b_i) f(|c_i|)
\]

**Table 5-3 KG Function \( h(a, b) \)**

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>A vector of current belief regarding the revenue for non-dominated choices – these are selected elements within ( E[Y^{d+1}] ).</td>
</tr>
<tr>
<td>( b )</td>
<td>A vector of the expected change in variance given we choose to measure price ( x ) for non-dominated choices – these are selected elements within ( \sigma(x) ).</td>
</tr>
<tr>
<td>( f(</td>
<td>c_i</td>
</tr>
</tbody>
</table>
5.3.2. Calculation of Correlated Belief Knowledge Gradient - Actual

Given the above discussion, we walk through – in the below section – the calculation for our case, in order to demonstrate intuitively why KG values can be effectively zero. In the below section, we explore the KG dip at $8.00. We graph the line graph (Figure 5-9) for our case below – we chose the graph from $Z = -2.0$ to $Z = 2.0$, where the majority of the probability weight lies, and only graph the top few lines for the sake of clarity. Each line represents how our expected revenue for a price would change. For instance, the light green line sketches how our expected revenue for price $8.50 will change depending on the Z-value. We see that in the figure below that even for relatively extreme Z-values ($Z = -1.5$ or $Z = 1.5$), our maximum choice is still $8.50. In other words, if we choose to measure $8.00$ for our next trial, the probability that it will change our choice of maximum is extremely low, hence earning it a low KG value.

**Figure 5-9 Line Plot for Measurement $p = 8.00$**

![Line Plot](image-url)
To view this from another perspective, we graph the updated beliefs of revenue below against prices in Figure 5-10 below. Each curve represents our new updated belief regarding the revenue curve (as a function of price) given a specific Z-value and that we choose to measure $p = 8.00$.

**Figure 5-10 Updated Belief Given Measurement $p = 8.00$**

![Graph showing updated beliefs and maximums given measurement $p = 8.00$]

The maximums of the new beliefs are circled in red; we again see that even for large values of Z, measuring $8.00$ does not change our choice of maximum. Since information is only valuable so far as it *changes our decision*, this makes sense that choosing to measure price of $8.00$ has little value, since it does not change our decision even for extreme values of Z.

In our original KG graph (Figure 5-6 and 5-7), we see that the KG value for points close to $8.00$ is substantial compared to that for $8.00$. Hence, we repeat similar analysis of line graph for $8.25$ and $7.75$ in Figure 5-11 below.
We see that compared to the graph for $8.00, the lines no longer look parallel, and for the measurement choice of $8.25, when Z-value gets high, the expected revenue for $8.75 (the purple line) seems to surpass that of the current maximum $8.50 (light green line). Alternatively, for the measurement choice of $7.75, when Z-value gets low, the expected revenue for $8.75 again seems to edge out our current maximum $8.50. We further confirm this with the updated belief and maximum graphs below for $7.75 and for $8.25:
Indeed, we see that if we choose to measure $7.75 and $Z=-2$, our new maximum will be at the price of $8.75; with other graphed Z-values, the maximum does not change. And if we choose to measure price $8.75 and $Z = 2$, we get again that $8.75 is the new maximum choice. Hence, we do see how the measurement choices of $7.75 and $8.25 have a higher probability of changing our decision than the measurement choice of $8.00, and hence have higher KG values. We repeat a similar experiment for price $3.50, which also has an effective KG value of zero (see previous figures Figure 5-6 and 5-7), and the results were comparable.

Recall that the measurement choice of $10.00 has the highest KG value (see Figure 5-6 and 5-7). We repeat a similar analysis for $10.00, which proves to be quite informative. Figure 5-13 contains the line graph for the measurement choice of $10.00,
and the potential updated belief and new maximum for the measurement choice of $10.00.

**Figure 5-13 Line Plot and Updated Beliefs Given Measurement p = $10.00**

We see that compared to the previous graphs, where the KG values were all relatively low, the line graph shows significant number of intersections, and our choice of the maximum is highly sensitive to Z-value. This is confirmed by the graph on the right, which demonstrates how our new maximum choice shifts significantly depending on our new observation. In other words, if we choose to measure price $10.00, our new observation provides significant chance of changing our next choice, and hence this information is valuable.

To further explore the behavior of KG under correlated beliefs, we graph a larger range of Z-value for the measurement choice of $10.00, and arrive at Figure 5-14 below:
Not only is the graph aesthetically fascinating, it also has some interesting properties from an optimal learning perspective. First, for look-up table correlated beliefs, as mentioned in Chapter 5 of Powell and Ryzhov, the slope of all the lines in a line graph can only be positive, since if we get a higher observation than expected, it should pull our whole expected revenue curve up. However, in parametric belief models, we can experience negative slopes – we see that if our observation for $10.00 is lower than expected (negative Z-value), our expected revenue for $7.00 (the solid teal line) actually becomes higher.

**5.3.3. Look-up Table vs. Parametric Beliefs: Some Intuitions**
The above section demonstrates that parametric belief models exhibit two interesting characteristic not found in typical look-up table model:

- Beliefs can have negative covariance
- KG value can be effectively zero

We seek to explain these differences in a relatively intuitive way in the following discussion. Overall, it is important to note that the below discussion regarding updating is actually independent of KG. Hence, these differences we discuss are inherent – and hence applicable – in all look-up table versus parametric belief models in Bayesian learning.

The parametric belief model differs from a look-up table model in that – unlike a look-up table’s expected revenue curve, which can be of any form – the expected revenue curve (our current belief) is forced into a specific functional form. For instance, if we compare how a look-up-table model versus a parametric belief that is linear in price (in form of $E[Revenue] = a + bp$) updates their belief, we can see why negative covariance is possible. Figure 5-15 below demonstrate some of the basic intuition:

**Figure 5-15 Updating Behavior of Parametric (Linear) vs. Look-Up Table**

![Expected Revenue Curve Comparison](image-url)
While the above figure does not use actual numbers, it does illustrate the key intuition that accounts for why there can be negative covariance. See that in the look-up table case, the much-lower-than-expected new observation only pull down the belief around its price point and does not have much effect on points further away. However, in the parametric model, since we require the belief to be a line, this new observation actually shifts the whole line, and we can see while this new observation is lower than expected, it actually increased our expected revenue for prices closer to the y-intercept. Hence, we see that negative covariance is possible under the parametric belief model. Obviously, this behavior is not exactly ideal; a lower-than-expected observation anywhere on the graph should lower our belief curve, not raise it. However, this is the price that one must pay when we attempt to gain efficiency – and many times feasibility of solving a problem – by using a parametric model; it allows us to keep less information from run-to-run, but we do have to accept some not-ideal behaviors.

A similar method can be employed to demonstrate the intuition for why KG value of effectively zero is possible for a parametric model, while they would not exist in a look-up table setting. Recall the line graphs in Figure 5-8, where the KG value is effectively zero when regardless of the observation, the maximum choice still remain the same. This can similarly be explained by parametric model’s requirement that the belief be a strict functional form. We compare a look-up table belief against a parametric belief model that is quadratic in price (in form of $E[Revenue] = a + bp + cp^2$). Figure 5-16 below demonstrate this intuition.
We see that for the case of look-up table, if we get a much-lower-than-expected observation, it affects only the points around it, while the old observations are still “propping up” the curve, so our maximum choice will change given this new observation. However, in a parametric belief model that forces the belief to be a quadratic function, such m-shaped curve is not possible; hence, the curve merely shifts down, and we get that the maximum choice remains the same; hence, the new measurement does not provide valuable information under the parametric belief model; when the behavior gets so extreme that for relatively rare Z-values the decision does not change, we get a KG value that is effectively zero.

While we can name these points “dead zones”, “fulcrum points” might be more appropriate, borrowing the idea from physics. These points are not necessarily the maximum, but are points where any relatively probable measurement will not change maximum choice. The rationale for naming them “fulcrum point” is demonstrated in the following figure. We picture the expected revenue curve of our parametric belief model as a beam balancing on a fulcrum point – not unlike a seesaw – and demonstrate the effect of new forces (representing new measurements) on the system:
We see that if an upward force is applied at the fulcrum point, the whole beam shifts up, and while the new belief is not the same as the old one, our choice of the maximum would not change. However, if we apply an upward force at the edge of the beam, we see significant shift in the curve, and our new maximum choice would change. We see that values at the edge have a much stronger potential to change the curve than a point near the fulcrum. This diagram also provides some intuition as to why KG tends to test edge value in a parametric belief model. A direct corollary of this statement is that KG is extremely powerful at identifying these “fulcrum points” given the current state of belief.

**5.3.4. Finding “Fulcrum Points”**

The next natural question to ask – given that this behavior is inherent in all Bayesian learning with parametric model and is independent of KG – is how one can determine which points are the “fulcrum points” without calculating the KG value each time? We recall that near fulcrum points, the KG value is effectively zero: this happens when any reasonable observation given that price point will not change our choice of
maximum in the updated belief, which implies that measuring that price yield little new valuable information. We hypothesize that in our exploration above, $8.00 is not the actual “fulcrum point,” but only close to it, since its KG value is not completely zero. The actual “fulcrum point” theoretically should yield a KG value of zero. We now seek to derive theoretically where this can happen.

Based on our discussion of the line graphs and how it links to new maximum, the actual fulcrum point should be at the price where for any possible observation of the measurement price, the maximum choice would not change in the updated belief. Translating this back to the line graph discussion above, this condition can only occur when the lines in line graph are exactly parallel, and hence that the upper envelope would just be the line representing our current maximum. This implies that the slope of all these lines give our measurement should be identical. In turn, this necessitates that elements in the vector below has to be identical at the x that constitutes the fulcrum point:

$$\sigma(x) = \frac{\Sigma^d e_x}{\sqrt{\Sigma^d_{xx} + \lambda^W}} = \frac{X(\Sigma^d \theta)(X)^T}{\sqrt{\Sigma^d_{xx} + \lambda^W}}$$

Since the square root below is simply a scalar, requiring all the elements of $\sigma(x)$ implies that the $x^{th}$ volume of $\Sigma^d$ must be filled with identical elements. To derive the point more specifically, we try out a linear case below, trying out three possible measurement locations, one of which is a variable representing the location of the fulcrum point we are trying to solve for; note that the covariance matrix is a symmetric matrix.

$$\sigma(x) \sqrt{\Sigma_{xx}^d + \lambda^W}$$

$$= X(\Sigma^d \theta)(X)^T e_x$$
We want to solve for $x_f$ such that these three equations equal to each other. Taking out identical terms in these three expressions, we get the following system of equation to solve:

$$\begin{bmatrix}
    c_{11} + c_{12}x_1 + x_f(c_{11} + c_{22}x_1) \\
    c_{11} + c_{12}x_2 + x_f(c_{11} + c_{22}x_2) \\
    c_{11} + c_{12}x_f + x_f(c_{11} + c_{22}x_f)
\end{bmatrix} = \begin{bmatrix} k \\ k \\ k \end{bmatrix}$$

We see an easy solution to this, with $k = 0$.

$$x_f = -\frac{c_{12}}{c_{22}}$$

We see that this solution is independent of other prices $x_1$ and $x_2$, and hence is independent of how we discretize our price spectrum. The quadratic case is much more difficult to solve by hand, but we believe that a similar approach can be taken.
Chapter 6: Policy Implications

While the model we used for this project is not completely realistic, and hence cannot give us a numerical estimate for how much demand response will reduce the cost for utility companies, it does provide some general qualitative results that should be of interest to policy makers. In the sections below, we first explore the qualitative results of this simulation with regards to demand response, and then proceed to discuss the two main challenges in implementing this type of demand response programs, namely participation by customers and regulation of utilities as monopolies. We conclude this section not with directives for policy change, but rather with key questions that still needs to be addressed, both from a legal and a moral standpoint. In order to provide more concrete discussion, we have drawn examples of demand response programs and current
regulatory frameworks from Illinois and California, which provides a relative abundance of information and data.

6.1. Impact of Demand Response

In running the simulation, we tested – in addition to our learning policies – a control policy that assumes there is no demand response and that electricity is charged on a cost-based process rather than market-based method; this is translated to charging the minimal price on our spectrum ($10.00). The previous result section discussed the profit difference of these policies, which takes the perspective of a utility company. However, from the perspective of a policy-maker contemplating the implementation of demand response, we are more concerned with whether demand response has achieved its stated goal – to redistribute demand in such a way as to increase reliability of the system. To that end, we also kept track, for each hour and each day, the percentage of garages that are over-capacity under each policy. Figure 6-1 below outlines the result of each of the simulation over 40 trials.
On the horizontal axis, \textit{ctrl} means the control case policy, where we simply charge the minimum price at all times, \textit{nl} is the no-learning case (a pure exploitation policy without updating of belief), \textit{exp, kg, ie, exc, boltz} means pure exploitation, knowledge gradient, interval estimation, excitation, and Boltmann exploration policies, respectively. Please see Chapter 3 for more detailed explanation on these policies.

The error bars above represent 95% confidence intervals regarding our estimate (namely, 1.96 times the standard deviation). We clearly see that comparing to the control policy, the rest of the policy all perform better or similar to the control policy. However, we see that once we have a clearer understanding of the demand curve (such as under interval estimation or knowledge gradient policy where beliefs eventually come closer to the truth), demand response has an even more powerful impact, and the difference, from the non-overlap of the confidence estimates, is statistically significant. The graph above reveals that demand response does make a difference, but only makes significant difference when one has a good understanding of the demand curve.
This result suggests that demand response can have significant potential to reduce congestion areas in the grid, potentially cutting the probability of overcapacity by half. This result parallels Federal Energy Regulatory Commission’s (FERC) estimates that dynamic pricing strategies has potential for 14-20% reduction in peak demand. While the FERC study focused on time-of-use demand response policies, our study further suggest that similar concept can be effectively applied to not just time, but to space as well, producing a location-of-use demand response policy. As the introduction lays out, switching to such demand response is much more economically efficient, because it aligns the demand with the actual cost it takes to supply electricity to the location of use, whether the cost is driven by price fluctuations in the wholesale market or grid failures within a local distribution network.

The qualitative result of this study is not restricted to Queens – similar concepts can be applied to any large city, such as San Francisco and Boston, that may face the dual-pronged challenge of the increasing number of electric vehicles and an aging grid.

Given that this location-of-use demand response policy has high potential to reduce impact on the distribution network, we next examine the larger context in which it will have to be implemented, specifically how past demand response program has met the challenge of customer participation and regulatory hurdles.

6.2. Customer Participation

At its core, demand response is about using price as a way to deliver information and to incentivize behavior change. This is a two-step process: the customers must be

---

34 Federal Energy Regulation Commission, National Action Plan on Demand Response, 5
able to receive the info and be willing to act on it. Currently, the majority of customers who participate in demand response programs are large commercial users, and there has been very little penetration of demand response into the typical consumer market; indeed, as FERC study found, “customers who participate in these programs are primarily in the commercial and industrial sectors and represent a relatively small fraction of all retail customers.”\(^{35}\)

That does not mean, however, that relatively successful demand response programs have not been found in residential space. A cursory search yields that California’s Pacific Gas and Electric (PG&E) and Illinois’ Commonwealth Edison (ComEd) both have implemented some sort of demand response program for its residential customers, and these programs both have been quite successful. PG&E has a “SmartRate” plan where customers’ regular rate is discounted, but there will be up to 15 days annually designated “SmartDays,” where the rate for the electricity will be much higher; these days are announced at least 24 hours in advance to allow for the consumer to adjust their behavior. This program has been quite successful, resulting in reduction of 16.9% in peak energy demand from customers who participated, and 81% of participants in 2010 are “very satisfied,” and especially low-income participants, of whom 91% were “very satisfied” with the program.\(^{36}\) ComEd’s program in Illinois installs a smart meter at the customer’s home that allows the utility to turn off customer’s air condition and refrigerator when the load on the grid becomes exceedingly high.\(^{37}\)

\(^{35}\) Ibid., 25
Despite the fact that demand response both benefits the utility and the consumers, and is more economically efficient for the whole market, participation rate has remained low. PG&E has enrolled roughly 0.15% of its customers, while ComEd’s program has only reached 0.06% of its customers.  

A study on the participation rate of these demand response programs revealed that one of the major difficulty is “getting [customers] to commit to 12 months of uncertain electricity costs.”

This brief survey of current demand response program seems to suggest that demand response program is beneficial to both the customer and the utility, but that overcoming the hurdle of uncertain price is difficult in a residential setting. The pricing scheme developed in this project specifically does not suffer from this type of difficulty, since the customers would know the price before they enter a garage and can change their behavior before committing. Furthermore, if this pricing scheme is implemented in all the garages, then customers would have to subscribe to demand response, since there would be no other options.

6.3. Regulatory Framework

While the pricing scheme proposed here overcomes the critical challenge of participation faced by other demand response programs, there are still significant regulatory hurdles to such a pricing scheme. In the below discussion, we first examine the larger context of regulation within the electricity market, and then proceed to analyze the specific cases of past demand response programs and their regulatory structure.

---

39 Ibid., 265
6.3.1. Regulation in the Electric Market

Electricity markets, unlike most markets for consumer goods, are regulated heavily, and there are strict limitations on what prices can be charged and when. Indeed, unlike a Coca-Cola or Honda, power companies cannot invest in new major projects or change the prices they charge without regulatory permission. The fact that utilities are natural monopolies and that electricity is a necessary good for most people drives the necessity of regulation; indeed, the Department of Energy states, “regulation of utilities is based on the inherent risk that a single monopoly supplier will overcharge consumers due to the lack of competition and high demand.”\(^{40}\) The power for the government to oversee the electricity market derives from the Public Utilities Holding Company Act of 1935, which was a response to exactly this problem of large cross-state utility companies charging extremely high prices for electricity due to their monopoly status.\(^{41}\) Indeed, we see that PG&E had to constantly cooperate with California Public Utilities Commission (CPUC) to enact the “SmartRate” program mentioned above, and CPUC passed a close to 100-page document prescribing the details of how PG&E can implement demand response pricing for its customers.\(^{42}\)

Before examining the specific regulation of pricing for demand response, however, we first survey the regulatory scene of the market as a whole, identifying which government players are paired with which level of the electricity market. Table 6-1

\(^{41}\) Ibid., 5.3
below presents a quick summary of the major player, both industry and government, at
each level of the market.

**Table 6-1 Major Players in Wholesale vs. Retail Markets**

<table>
<thead>
<tr>
<th>Major Roles</th>
<th>Wholesale Market</th>
<th>Retail Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation</td>
<td>• Federal Energy Regulatory Commission</td>
<td>• State Government</td>
</tr>
<tr>
<td></td>
<td>• National Energy Reliability Corporation</td>
<td>• Local Government (city, county)</td>
</tr>
<tr>
<td>Supplier / Producer</td>
<td>• Generator (power plants)</td>
<td>• Electricity Retailer (utilities)</td>
</tr>
<tr>
<td></td>
<td>• Transmission Services</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ancillary Services</td>
<td></td>
</tr>
<tr>
<td>Buyer</td>
<td>• Electricity Retailer (utilities)</td>
<td>• Customer (residential, business, government)</td>
</tr>
<tr>
<td></td>
<td>• Large Retail Customers (industrial)</td>
<td></td>
</tr>
</tbody>
</table>

We roughly divide up the electricity market into wholesale and retail; wholesale is concerned with production of electricity and the delivery to the local grid, while the retail is concerned with distributing it to the actual consumers of electricity; wholesale markets tend to be across multiple states, while the retail market tends to be much more local – on a state or even county level.

The wholesale market is overseen by the FERC, the over-arching national regulatory body that "regulates the transmission and wholesale sales of electricity"
in interstate commerce,” and is also the oversight for major mergers and acquisition in the electricity industry. On a more day-to-day level, the federal government has given the responsibility to the North America Electricity Reliability Corporation, which is charged with “enforce[ment of] reliability standards for the bulk-power system.” Major industry players in the wholesale market includes the generator of electricity, such as power plants, transmission service providers that brings electricity from generator to distributor, and finally ancillary service providers that help bring stability to the system. The buyers of electricity at this level are retailers that sell electricity to consumers; occasionally, large industrial consumers who have the expertise and the bulk to participate in the much more volatile wholesale market are also buyers in this space. Independent system operators (ISO) or Regional Transmission Organization (RTO) are also involved in this market; these players organize the regional wholesale market to connect suppliers with buyers, and oversees the minute-to-minute reliability of their regional grid.

State and local government, on the other hand, oversees the retail market, which is the market most consumers interact with. The state government sets the rate at which utilities can charge the customer, while the local government has say over the construction of new plants or lines over its territory. Section 6.3.2 below takes up the discussion of price regulation in more detail. On the side of the industry, we have utilities and retailer of electricity as the supplier of goods in this

market, and the end consumer as the buyers; this is the typical relationship between a local utility, such as ConEd, and an end user, such as a resident in New York City.

Note that this is a simplification of the market, and that the actual market can be much more complex; indeed, many utilities are quite vertically integrated; for instance, Exelon Power Company owns significant generating capacity, and its subsidiaries participate as local utilities in the retail market in Maryland, Illinois (including Chicago), and Pennsylvania.45

6.3.2. Price-Setting Process for Local Utilities

The price that a customer sees on an electricity bill is not arbitrarily determined by the utilities company, but rather through a complicated and lengthy review through the state regulatory commission; indeed, a survey of regulatory regime in each state reveals that commissions in each state has a highly legal procedure for determining electricity rates for different type of customers, with well-defined filing, witness, hearing, and appeals processes.46 More specifically, these cases of setting prices for electricity are called tariff cases, where “tariff” is defined as “a commission’s approved conditions, terms, and prices of utility services.”47 In most tariff cases, two major types of decisions are made – first rate level, or what level of revenue should the utility receive in order to make a “reasonable return,” and rate design, or how that cost should be distributed among different classes of customers. The determination of rate level is steeped in financial considerations – the commission first determine the

47 Ibid., 31
basic revenue requirement for the utility to make a reasonable profit, and then readjust the revenue so that the utility can deliver a “reasonable rate of return” to its stakeholders given the capital structure of the firm.48

Of more interest to us is how this cost is allocated. The commission first defines the classes of customers, typically residential versus commercial, though they can be sorted by other aspect as well; for instance, government entities often get its own class. Then, the commission then defines what types of rate the utility can charge under what circumstances to each class of customers, and this process is not just based on economics; indeed, as the Regulatory Assistance Project summarized, “rate setting, and especially allocation decisions, can be partly judgmental and partly political, not just technical.”49

6.3.3. Current Dynamic Pricing Programs

Given the complicated regulatory framework, it is useful to examine what types of dynamic pricing program have been approved in the past, to establish the “frontier” of regulation in regards to dynamic pricing. Current dynamic pricing program can be roughly categorized into three types depending on how far in advance the rate is “set”: time of use pricing, critical peak pricing, and real-time pricing.

- **Time of Use Pricing (TOU)** – rates for electricity depends on time blocks, and these rates are announced significantly in advance, so is based on

48 Ibid., 38
49 Ibid., 50
historical conditions rather than current conditions. Most large commercial users of electricity have moved to TOU pricing; Illinois required this for large industrial users, and in California, CPUC notes that “virtually all large customers had moved to time-of-use (TOU) rates.”\textsuperscript{50}

- **Critical Peak Pricing (CPP)** – this is similar to TOU, but the rates are announced much shorter in advance (usually a day or two) as to more properly reflect current conditions rather than historical data; indeed, CUPC differentiate this from TOU by the fact that while “the time and duration of the price increase are predetermined,… the days are not predetermined.”\textsuperscript{51} PG&E’s “Smart Rate” program is one example of CPP pricing.

- **Real-Time Pricing (RTP)** – this is the ideal form of demand response, defined by CPUC as “a dynamic rate that allows prices to be adjusted frequently, typically on an hourly basis, to reflect real-time system conditions.”\textsuperscript{52} This allows the price of electricity to reflect the current condition almost completely, and hence is the most efficient, at least from view of an economist. Traction amongst consumers, however, can be quite difficult due to uncertainty regarding electricity prices, especially because the consumers may be away from home and cannot change behavior fast enough to avoid pricing spikes. ComEd’s current demand

\textsuperscript{50} California Public Utilities Commission, *Decision Adopting Dynamic Pricing Timetable and Rate Design Guidance for Pacific Gas and Electric Company*, 5
\textsuperscript{51} Ibid., 6
\textsuperscript{52} Ibid., 7
response is a specific instance of RTP pricing, and as we have seen, the participation rate has been quite low. The demand response policy we have modeled in this thesis is a real-time pricing strategy.

6.4. Challenges of Implementing Location-of-Use Demand Response for Electric Vehicles

The above analysis reveals that a variety of demand response programs have been implemented, and that the state commissions are the major bodies through which these types of programs have to gain approval. Nevertheless, refueling of EV's represents a new market, and significant uncertainty lies ahead. For the regulatory commission facing this decision, there are several key questions that still needs to be addressed: the type of regulation EV charging market should be under, whether utilities can enter this type of market, and additional legal and ethical concerns regarding such location-of-use demand response programs.

The first question to be considered is how this EV fueling market should be regulated, if at all. Indeed, one can argue that EV fueling stations are equivalent to gas stations, which receive little regulation, but one can also argue that EV fueling stations are more like residential electricity market, since it relies so heavily on the grid. Since most parking garages would not have the expertise of installing and running electricity charging spots, we predict a cooperation between the utilities and the garages, where the utilities will be involved in helping install these charging ports and then charging the drivers directly for the refueling of their car. The garage gains from being able to attract EV customers, and the utility gains from being able to serve more
customers. As a hybrid market of sort, it is still unclear what type of regulation will this market be subject under, and how much oversight the state commissions should have.

Secondly, regardless of whether this type of market will be under regulation, the utilities companies will still be heavily regulated, and their entry into such a market will be called into question, since the state commission does oversee key investments of utilities to ensure that they align with the public interest, and that utilities are operating to the best of their capabilities and not relying on the revenue “safety net” in making risky or unprofitable investments. And even assuming that the utilities can enter this market, there is still the question that all demand response programs must face – how much flexibility can the company have in changing the prices? Indeed, the simulation in this thesis suggest that rapid flexibility can be quite beneficial from an economic standpoint, but the commission has to balance political and ethical concerns as well.

On the ethical front, it can be a concern if a location is constantly congested. While economically it makes sense to charge a higher price for an area that is often congested, if such an area coincides with a low-income neighborhood, we can see from a Rawlsian perspective how such public policy cannot be justified. And overall, since utilities are a monopoly, thorough research needs to be done to ensure that the utility is charging prices more to decrease the impact on the grid (while still earning a reasonable return), not to maximize profits.
Given the challenge of PHEV’s expected impact on the electric grid, including local distribution networks, we proposed location-of-use demand response policy as a potential solution. More specifically, we study the challenge of learning an unknown demand curve. We utilized a simulated model of truth order to study the problem of optimal learning from the perspective of an electric utility and public policy makers, and suggested that such type of demand response can be highly effective if utilities are given the flexibility to change prices and have a good understanding of – or at least given a chance to learn – the demand curve. The question of whether this type of flexibility should be given to utilities fall under the prerogative of state energy commissions, which must balance economic, political,
and ethical considerations in chiseling out a new regulatory framework for the electric vehicle refueling market. We also take a side trip to explore the existence of “fulcrum points” within the parametric Bayesian learning model.

While the results of this project provide great qualitative results and intuitions for why such results occur, the thesis is limited in scope, and many potential improvements can be made upon this project.

First, we calculated the electricity capacity of each garage independently, ignoring potential interactions between the loads from each garage on the grid. A more wholesome simulation would actually link each garage to a simulated form of Queens grid, and use actual load data from Queens to more closely align the simulation with reality.

Secondly, we did not fully explore the underlying game-theoretic nature of this problem. A natural extension of this problem would be to simulate or derive how long it would take for this new market to come to equilibrium, and what price point or behavior this equilibrium will entail.

Lastly, while we set up a way to solve for the “fulcrum points” in the parametric Bayesian learning model, and demonstrated it for the linear case, it would be quite valuable to both practitioners and theorists alike to be able to derive a theoretical closed-form formula for any degree of polynomial parametric belief.
References


DeForest, Nickolas, Jamie Funk, Adam Lorimer, and Boaz Ur. *Impact of Widespread Electric Vehicle Adoption on the Electrical Utility Business – Threats and Opportunities* 2009.


Powell, Warren. *Approximate Dynamic Programming*
Powell, Warren and Ilya Ryzhov. *Optimal Learning*  


