Submitted to *Transportation Science*
manuscript TS-2014-0299

Optimal Policies for the Management of an Electric Vehicle Battery Swap Station

Rebecca S. Widrick, Sarah G. Nurre, Matthew J. Robbins
Air Force Institute of Technology, Department of Operational Sciences, Sarah.Nurre@gmail.com

Optimizing operations at electric vehicle (EV) battery swap stations is internally motivated by the movement to make transportation cleaner and more efficient. An EV swap station allows EV owners to quickly exchange their depleted battery for a fully charged battery. We introduce the EV-Swap Station Management Problem (EV-SSMP), which models battery charging and discharging operations at an EV swap station facing nonstationary, stochastic demand for battery swaps, nonstationary prices for charging depleted batteries, and nonstationary prices for discharging fully charged batteries. Discharging through vehicle-to-grid is beneficial for aiding power load balancing. The objective of the EV-SSMP is to determine the optimal policy for charging and discharging batteries that maximizes expected total profit over a fixed time horizon. The EV-SSMP is formulated as a finite-horizon, discrete-time Markov decision problem and an optimal policy is found using dynamic programming. We derive structural properties, to include sufficiency conditions that ensure the existence of a monotone optimal policy. Utilizing available demand and electricity pricing data, we design and conduct two computational experiments to obtain policy insights regarding the management of EV swap stations. We compare the optimal policy to two benchmark policies that are easily implementable by EV swap station managers. Policy insights include the relationship between the minimum battery level and the number of EVs in a local service area, the pricing incentive necessary to encourage effective discharge behavior, and the viability of EV swap stations under many conditions.

Key words: Green logistics; Markov Decision Processes; Monotone Policy; Electric Vehicles
History:

1. Introduction

Optimizing operations at electric vehicle (EV) battery swap stations is internally motivated by the movement to make transportation cleaner and more efficient. The U.S. Energy Secretary, Ernest Moniz announced a $50 million budget in January 2014 for research of vehicle technologies, which will also aid the initiative launched in March 2012 to make plug-in electric vehicles more convenient and affordable over the next 10 years (U.S. Department of Energy 2014). We approach this research

initiative by considering the optimal management of EV battery swap stations. An EV battery swap station allows the EV owner to exchange their depleted battery for a fully charged one. By implementing swap stations, not only are EV owners offered the convenience to swap their battery, but there is the opportunity to control battery charging and reduce the negative effect of increased demand for electricity on the power grid (Clement-Nyns et al. (2010), Bingliang et al. (2012)) and reduce the difference between high-peak and low-peak energy prices (Eyer and Corey 2010).

The concept of battery swap stations for electric vehicles was initially developed by the Israeli company Better Place, who financially collapsed in May 2013 (Pearson and Stub 2013). Despite Better Place's collapse, it is still of great interest to examine such swap stations as the manufacturing of plug-in hybrid electric vehicles (PHEVs) and EVs is on the rise and the motivation to switch from gasoline to battery power remains undiminished. According to the U.S. Department of Energy (2014), nearly 100,000 EVs were purchased by Americans in 2013, almost twice as many as in 2012.

One of the leading electric vehicle manufacturers, Tesla, first gained worldwide attention when it released the first ever mass produced electric powered sports car in 2010 (Abreu 2010). The Tesla Model S (sedan) is the current model available for purchase with two battery options and is marked at $71,070 for the 60 kWh battery option, $81,070 for the 85 kWh battery option, and $94,570 for the 85 kWh performance model. The Model X (crossover) has recently been unveiled and is currently available for reservation with delivery expected in Fall 2015 (Tesla motors 2014c). A third model is said to be released in 2017 at a cost of $35,000 by the Tesla founder and CEO, Elon Musk (Fowler 2014). It will be called the Model 3 and will be a direct rival of the current BMW 3 Series electric car. The rolling out of electric vehicles to the market is also occurring for many other vehicle manufacturers. Honda, BMW, Chevrolet, Ford, Nissan, Cadillac, Fiat, Mercedes, Mitsubishi, SMART, Volkswagon, Kia, and Toyota all carry at least one electric vehicle and can cost between $23,800 for the Mitsubishi i-MiEV to $137,000 for the 2014 BMW i8 (Plug-In Cars 2014).

In addition to being one of the leading electric car manufacturers, Tesla is also the frontrunner when it comes to charging stations. There are currently 129 Tesla supercharge stations in North America, 95 in Europe, and 36 in Asia (Tesla motors 2015). Electric car owners can plug in their car at a supercharge station and receive 120 kW of charge in just 30 minutes at no cost to the consumer. This provides 170 miles of travel for the Model S 85 kWh battery option. While this is a great option for EV owners, it still requires a wait time while the battery is charging and plug-ins may get congested as the number of EVs purchased continues to increase. Battery swap stations provide a fast and convenient way to drive away with a fully charged battery. Tesla presented the

idea of swap stations in June 2013, but they have not yet come to market (Tesla motors 2014a). However, Tesla is currently testing a pilot swap station program in California (Tesla motors 2014b).

Widely available battery swap stations will help the movement launched in March 2012 by the U.S. Department of Energy (2014) to make plug-in electric vehicles more convenient and affordable, as well as help control battery charging to avoid loss of power and power quality that can be incurred when batteries are charged during high peak demand for electricity (Clement-Nyns et al. 2010). An ancillary benefit of a swap station is the ability to coordinate discharging back to the power grid through vehicle-to-grid (V2G) technology (Sioshansi and Denholm 2010). When the charging and discharging of batteries is properly coordinated with the power grid, load balancing can occur (see Peng et al. (2012), Wang et al. (2011), and Göransson et al. (2010)).

With the significant impact swap stations can have on the growing market for battery powered vehicles, it is valuable to develop a model that optimizes the operations at a swap station. As such, we wish to model the system to reflect uncertainty of battery swap demand and nonstationary charging costs to gain realistic results that are robust to the stochasticity of the system. Thus, we consider the EV-Swap Station Management Problem (EV-SSMP). To model the EV-SSMP we develop a Markov decision process model (Puterman 2005). Markov decision processes characterize problems with discrete time sequential decision making under uncertainty and can be solved using dynamic programming. They can be modeled using finite or infinite horizons. Infinite horizon models provide for the determination of a stationary optimal policy, meaning that the optimal action is state dependent and not time dependent. Nonstationary Markov decision processes relax the assumption that problem data does not change with time and are in general unsolvable using infinite horizon models due to infinite data requirements (Ghate and Smith 2013). We consider a finite horizon model because our problem data is highly variable with respect to time. The nonstationary variable properties include mean demand for battery swaps, charging price for batteries, and revenue from discharging batteries back to the power grid. In a sequential decision making model, the state of the system is observed at a certain point in time and an action is taken. The action results in an immediate reward to the decision maker and the system transitions to a new state according to a probability distribution determined by the chosen action.

The Markov decision process for the EV-SSMP is characterized by the following: (1) decision epochs are a consistent time unit at which a swap station manager needs to determine the number of batteries to charge or discharge; (2) the state of the system is the total number of batteries that are fully charged, where the state of any given battery is either fully charged or depleted; (3) the action space is defined as one dimensional, where the decision maker chooses the total number of batteries to charge or discharge; (4) the reward function is comprised of revenue from battery swaps, revenue from discharging batteries back to the power grid, and cost from charging batteries;

and (5) transition probabilities are determined by customer demand for battery swaps (which we assume follows a discrete distribution), the current state, and the chosen action.

The objective in solving our Markov decision problem (MDP) is to determine a policy that maximizes the expected total reward criterion. A policy consists of decision rules that indicate to the decision maker an action to take in a given state at a given point in time. For the EV-SSMP, a decision rule specifies the number of batteries to charge or discharge at a given point in time given the current inventory of fully charged batteries. We prove that when the demand for swaps follows a discrete nonincreasing distribution that a monotone nonincreasing policy is optimal. The optimal policy is found using the backward induction algorithm (Puterman 2005)for the general case and the monotone backward induction algorithm (Puterman 2005) for the case when demand is governed by a nonincreasing discrete distribution. We compare the optimal policy to two benchmark policies that are easy to implement at the swap station. In the first benchmark policy, denoted the stationary benchmark policy, we assume the swap station maintains a single target inventory level of fully charged batteries regardless of time of day and day of week. In the second benchmark policy, denoted the dynamic benchmark policy, we assume the swap station maintains a distinct target inventory level for each time period (which captures time of day and day of week information). Each target level is based on the number of batteries at the swap station, charging costs, and the mean demand. The action for each policy is calculated by taking the difference between the current state of full batteries and the target level. If the swap station has more fully charged batteries than the desired level, it will discharge down to the target. If the swap station has less fully charged batteries than the desired level, it will charge up to the target.

We computationally test the optimal solution methods and two benchmark policies to gain insight regarding the optimal operations and policies that should be implemented at an EV swap station. We perform two Latin hypercube designed experiments. The first experiment is conducted to gain overall information for various parameter inputs for the swap station. Specifically, we examine the external factors, including the effects of uncertain demand, the incentive that should be given by the power company for discharging, and the seasonal charging cost variations. The second experiment is conducted to gain insight concerning the controllable internal parameters at a swap station (e.g., the number of batteries and swap price) in relationship to the number of EV swaps and power prices. Further, from the results of the second experiment we conclude that the dynamic benchmark policy outperforms the stationary benchmark policy; however, both exhibit the favorable characteristic of ease of implementation.

Growing interest in electric powered vehicles has led to extensive research on the topic in both industry and academia. Herein, we discuss relevant literature pertaining to the EV swap station

application and proposed solution approach. To our knowledge, no past research utilizes an inventory control MDP to model the operations of an EV swap station to decide the number of batteries to charge and discharge when factoring in stochastic demand, nonstationary charging costs, and nonstationary revenue from discharging back to the power grid.

Other studies have been conducted looking to optimize operations at EV swap stations in different contexts. The most similar to our study is the work of Worley and Klabjan (2011) who propose a dynamic programming model that seeks to determine the number of batteries to purchase and charge over time while minimizing the total cost comprised of purchase price, charging cost, opportunity cost of unused batteries, and a penalty for unmet demand. Therefore, the actions determined by their model are motivated by a different set of costs and do not include the ability to discharge back to the grid using V2G. In comparison to the exact solution method we propose, they approximate solutions by fitting the value function with a separable piecewise linear function. Nurre et al. (2014) do consider the option to both charge and discharge at a swap station, however they make the assumption that demand for exchanges is known over all time periods. Moreover, they solve their problem utilizing a mixed integer programming formulation. Using an adequacy model and Monte Carlo simulation, Zhang et al. (2012) determine the adequate number of batteries to set for swapping over time when batteries can be used for both swapping and discharging. However, they do not capture the charging actions that would need to take place at a swap station.

Prior research also examines the specific infrastructure of charging and swapping stations in an area. Pan et al. (2010) consider a two-stage stochastic program that seeks to locate stations in the first stage and then once a demand scenario is realized at each swap station, the second stage determines the allocation of batteries for swapping and discharging back to the grid. Their model does not consider the dynamics and changing actions over time that a swap station manager would need to make. Using robust optimization, Mak et al. (2013) decide where to locate swap stations when the information regarding adoption rate of PHEVs is limited. They aid in determining a deployment strategy for locating swap stations as the success of each swap station is sensitive due to this limited information. Morrow et al. (2008) analyze the infrastructure requirements for the charging of PHEVs in residential settings as well as in commercial settings. They report that the availability of charging infrastructure allows the vehicles to require reduced energy storage capability and thus reduces the overall cost of purchasing the vehicles. Transportation system costs can also be reduced by providing rich charging infrastructure rather than using larger batteries to compensate for lesser infrastructure. Tang et al. (2012) examine optimizing the allocation of physical infrastructure space at a swap station between batteries and photovoltaic power generation capabilities.

A complementary thread of research is the use of EVs or other energy storage devices to solely balance the fluctuations occurring from the demand for power and other integrated highly variable renewables such as wind and solar energy. These problems are often solved via a similar methodology of dynamic programming. Sioshansi et al. (2014) examine energy storage with the power grid and estimate the capacity value, a metric used to quantify a resources's impact on system reliability. Solving their model using dynamic programming, they show that capacity values are sensitive to energy prices with variability of up to 40%. Using approximate dynamic programming, Salas and Powell (2013) consider multiple energy sources (e.g., pumped-hydro, batteries, flywheels) and determine near optimal time dependent control policies. Using an energy storage problem that seeks to determine the optimal flow of energy from the power grid to a battery and from the battery to demand over time, Scott et al. (2014) test a range of approximate dynamic programming methods.

The MDP employed for solving the EV-SSMP is in the class of inventory control MDPs. Inventory control MDPs have been utilized to model a wide range of applications including supply chain management (Giannoccaro and Pontrandolfo 2002), supply chain management with disruptions (Lewis 2005), airline seat control (Zhang and Cooper 2005), paper manufacturing (Yin et al. 2002), and assemble-to-order systems (ElHafsi 2009).

Main Contributions. The main contributions of this work are as follows: (1) development of a Markov decision process model to determine the optimal number of batteries to charge and discharge at an EV swap station when factoring in stochastic, nonstationary swap demand, nonstationary charging costs, and nonstationary discharging revenues; (2) proving the existence of a nonincreasing monotone optimal policy structure when demand is governed by a discrete nonincreasing distribution; (3) implementation of two exact solution methods; (4) generation of two benchmark policies that are easy to implement by a swap station manager; and (5) analysis of the results from two designed experiments, which provide policy insights for the effective management of an EV swap station.

The remainder of this paper is organized as follows. In Section 2 we formally define our problem as an inventory control MDP to include decision epochs, state space, action sets, reward function, and transition probability function. We theoretically prove that the EV-SSMP contains a nonincreasing monotone structure, which motivates the optimal and two benchmark policy solution methods presented in Section 3. In Section 4, we computationally validate the proposed model and solution methods by conducting two designed experiments and analyzing the results to arrive at policy insights. We conclude in Section 5 and provide opportunities for future study.

2. Problem Statement

We seek to solve the EV-SSMP by determining the optimal number of batteries to charge and discharge over time. Modeling this problem as a Markov decision problem (MDP), we factor in stochastic, nonstationary demand, nonstationary charging costs, and nonstationary revenue from discharging. We consider a finite horizon, single product inventory control model because our problem data is highly variable with respect to time. The nonstationary characteristics of the EV-SSMP include demand for battery swaps, charging price for batteries, and revenue from discharging batteries back to the power grid. Motivating the decision that comprises the optimal policy is the maximization of profitability at a single swap station.

Within the MDP model, we define our state as the total number of batteries that are fully charged. We model the state of the batteries at a fundamental level where each battery is either fully charged or depleted. A solution where charging and discharging occur simultaneously can be equivalently represented as solely charging or solely discharging when the discharging revenue is less than or equal to the charging price. Thus, we model our system such that we never charge and discharge batteries simultaneously. If the discharging revenue is greater than the charging price, we make the simplifying assumption that the EV station solely charges or solely discharges at any point in time. We may discharge up to the minimum of the total number of batteries that are fully charged and the total number of plug-ins available. In this context, what we denote a plug-in is the physical entity at a swap station that connects a battery to the power grid, thereby allowing it to draw from the power grid (i.e., charge) or discharge using V2G. The total number of plug-ins or what we denote as the charging capacity is assumed to remain constant over time. Similarly, we may charge up to the total number of batteries that are in the depleted state, provided that our charging capacity is not exceeded. Thus, the total number of batteries at the swap station is constant over time.

We model the system such that batteries charged at time t become full in time $t+1$. For a time period of one hour, this charging capability is comparable to that of Tesla superchargers (Tesla motors 2015). Batteries that are discharged take one time period to deplete but are immediately unavailable for exchange. Only fully charged batteries are available for exchange or discharging. Furthermore, batteries that are fully charged are always swapped if available when demand arrives. The cost to charge and the revenue from discharging batteries is realized during the time period in which the decision is made. We do not permit backlogging of demand as we assume customers will not wait at the station if batteries are unavailable. We use the expected reward criterion to capture revenue from battery swaps, revenue from discharging batteries back to the power grid through V2G technology, and cost to charge batteries at the swap station. The event timing for the EV-SSMP is outlined in Figure 1. We mathematically characterize the MDP for the EV-SSMP using the following notation.

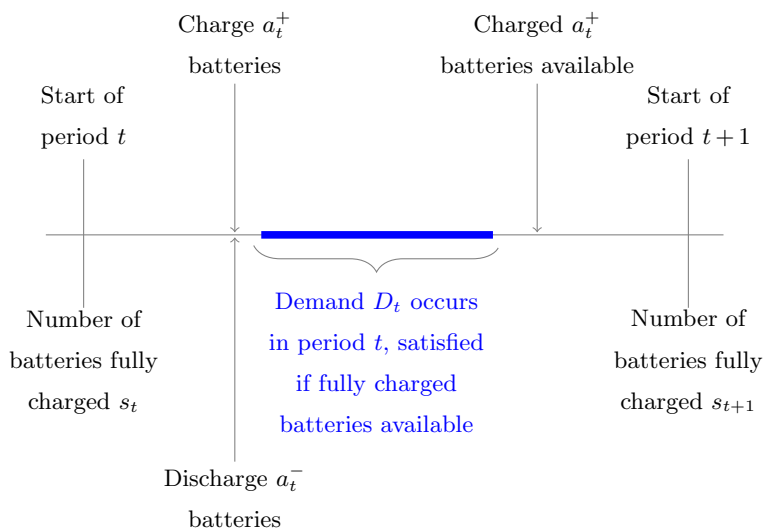


Figure 1 **Diagram outlining the timing of events for the EV-SSMP MDP model.**

1. The set of decision epochs[1], $T = \{1, \ldots, N-1\}$, $N < \infty$, indicates the discrete time periods in which a decision is made. As previously stated, we consider a finite time horizon due to nonstationary properties.
2. The state of the system at time t, $s_t \in S = \{0, 1, \ldots, M\}$ indicates the total number of batteries that are fully charged at decision epoch t, where M is defined as the total number of batteries at the swap station; thus, $M - s_t$ is the number of depleted batteries at time t.
3. The action at time t, $a_t \in A_{s_t} = \{\max(-s_t, -\Phi), \ldots, 0, \ldots, \min(M - s_t, \Phi)\}$, $\forall s_t \in S$ indicates the total number of batteries to charge or discharge at time t, where Φ is the charging capacity of the system. A negative action indicates the discharging of batteries and a positive action indicates the charging of batteries. For clarity in our model, we further define our action space. Let
$$a_t^+ = \begin{cases} a_t & \text{if } a_t \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$
and
$$a_t^- = \begin{cases} |a_t| & \text{if } a_t < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$
where a_t^+ is the number of batteries charged and a_t^- is the number of batteries discharged at time t. An assumption of the model is that a_t^+ and a_t^- cannot both be positive at any time t.
4. The immediate reward when action a_t is selected in state s_t at time t that leads to a transition to state s_{t+1} is the profitability of the system, given by
$$r_t(s_t, a_t, s_{t+1}) = \rho[s_t + a_t - s_{t+1}] - K_t a_t^+ + J_t a_t^- \tag{3}$$

[1] Decision epoch and time period will be used interchangeably throughout this paper.

for $t = 1, \ldots, N-1$, where $s_t + a_t - s_{t+1} = \min\{D_t, s_t - a_t^-\}$ is the number of batteries swapped at time t. Discrete random variable D_t represents the demand for battery swaps at time t, $s_t - a_t^-$ is the number of batteries available for exchange, ρ is the revenue per battery swap, K_t is the charging cost per battery at time t, and J_t is the revenue earned per battery discharged at time t. Specification of K_t and J_t captures the impacts of the nonstationary price for power over time. We calculate the terminal reward as potential swap revenue from fully charged batteries; thus, $r_N(s_N) = \rho s_N$.

5. The total number of batteries fully charged at decision epoch $t+1$ is directly impacted by the batteries charged, discharged, and exchanged during time period t by way of $s_{t+1} = s_t + a_t - \min\{D_t, s_t - a_t^-\}$. We define the probability of transitioning to state j at time $t+1$ from state s_t when action a_t is taken, denoted $p_t(j|s_t, a_t)$, by

$$p_t(j|s_t, a_t) = \begin{cases} 0 & \text{if } j > s_t + a_t \text{ or } j < a_t^+ \\ p_{s_t+a_t-j} & \text{if } a_t^+ < j \leq s_t + a_t \\ q_{s_t+a_t-j} & \text{if } j = a_t^+ \end{cases} \tag{4}$$

where $p_j = P(D_t = j)$ and $q_u = \sum_{j=u}^{\infty} p_j = P(D_t \geq u)$. For further clarification, $s_t + a_t - j$ indicates the number of fully charged batteries that are swapped in period t, and $s_t + a_t$ indicates the number of fully charged batteries on hand at the end of the period if none are swapped.

When the transition probability is zero, state j exceeds the number of fully charged batteries the swap station could possibly have on hand at the end of the period $(s_t + a_t)$ or state j is less than the number of batteries the swap station chooses to charge (a_t^+), which are not available for exchange until after demand is met in that period. In both cases there is a zero transition probability, as described in the first conditional of the probability function (4).

When the transition probability is $P(D_t = s_t + a_t - j)$, as in the second conditional of the probability function (4), state j is between the number of batteries the swap station charges and the number of batteries that could possibly be on hand at the end of the period. In this situation, the swap station has enough fully charged batteries to meet demand, hence the probability of transitioning to state j is calculated using the time dependent discrete distribution of demand. We have already established that j cannot fall below the number of batteries charged in that period, thus the lower bound on j is a_t^+.

When the transition probability is $P(D_t \geq s_t + a_t - j)$ where $j = a_t^+$, as in the last conditional of the probability function (4), demand for battery swaps meets or exceeds the supply of fully charged batteries available for swapping. In this situation, the station swaps all batteries on hand but acquires the charged batteries at the end of the period. The transition probability in this case is calculated using the cumulative probability that demand meets or exceeds the number of batteries available for swapping in period t.

To aid the reader, we illustrate the transition probability function using an example. Consider the case where there are 15 fully charged batteries and 35 depleted batteries in inventory (i.e., $s_t = 15$) and the swap station charges 5 of the depleted batteries (i.e., $a_t = a_t^+ = 5$). If no batteries are swapped the station will have a total of 20 fully charged batteries at the end of the period (i.e., $s_{t+1} = j = s_t + a_t = 20$). There is no possible way to have more than $s_t + a_t = 20$ fully charged batteries at the end of the period, thus there is a zero transition probability to a state greater than 20. At the beginning of the period there are $s_t + a_t^- = 15$ batteries available for exchange, thus if all fully charged batteries are swapped, the station still acquires the 5 batteries that were charged by the end of the period. Therefore, the transition probability to a state less than $a_t^+ = 5$ is zero. When $j = a_t^+ = 5$, the 15 batteries that were available at the beginning of the period must have been swapped, since the 5 charged batteries are acquired at the end of the period. The transition probability in this case is the probability that demand meets or exceeds $s_t + a_t - j = 15$ batteries, which is captured in the third conditional. Consider the case when the station has 7 batteries at the end of the period (i.e., $j = 7$, which is between a_t^+ and $s_t + a_t$). We know that 5 batteries were charged, leaving 2 remaining from the inventory in the previous period. Since the station started with 15 charged batteries, 13 of them must have been swapped. Thus, the transition probability to 7 batteries is the probability that demand for battery swaps was equal to $s_t + a_t - j = 15 + 5 - 7 = 13$.

Having specified the transition probability function $p_t(j \mid s_t, a_t)$, we are now able to express the immediate expected reward function in terms of the current state and action only (see Equation (5)), which is more desirable for subsequent calculations.

$$r_t(s_t, a_t) = \sum_{s_{t+1} \in S} \Big[p_t(s_{t+1} \mid s_t, a_t) \big(\rho[s_t + a_t - s_{t+1}] \big) \Big] - K_t a_t^+ + J_t a_t^- \tag{5}$$

We denote the decision rule functions, $d_t(s_t) : s_t \to A_{s_t}$, which indicate to the decision maker how to select an action $a_t \in A_{s_t}$ at a given decision epoch $t \in T$ when in state $s_t \in S$. Because our decision rules depend on the current state of the system and not the entire history of states, we consider Markovian decision rules (Puterman 2005). Furthermore, our decision rules prescribe a single specific action and not a probability distribution on the action set. Therefore our decision rules are deterministic. A policy π is a sequence of decision rules $(d_1^\pi(s_1), d_2^\pi(s_2), \ldots, d_{N-1}^\pi(s_{N-1}))$ that specify the decision rule to be used at all decision epochs.

The expected total reward of a policy π, when the initial state of the system is s_1, denoted $\upsilon_N^\pi(s_1)$ is given by

$$\upsilon_N^\pi(s_1) = \mathbb{E}_{s_1} \left[\sum_{t=1}^{N-1} r_t(s_t, a_t) + r_N(s_N) \right]. \tag{6}$$

We seek to determine the policy π^* with the maximum expected total reward. The optimal value function, $u_t^*(s_t)$, denotes the maximum over all policies of the expected total reward from decision

epoch t onward when the state at time t is s_t. We consider optimality equations, or Bellman equations, that correspond to our optimal value functions as a basis for determining the optimal policies. The optimality equations are given by

$$u_t(s_t) = \max_{a_t \in A_{s_t}} \left\{ r_t(s_t, a_t) + \sum_{j \in S} p_t(j|s_t, a_t) u_{t+1}(j) \right\} \tag{7}$$

for $t = 1, \ldots, N-1$ and $s_t \in S$. For $t = N$, we have $u_N(s_N) = r_N(s_N)$. The solution to the optimality equation at $t = 1$ gives the expected total reward for the entire time horizon.

Since the management of an EV swap station is perpetual, a discussion is warranted concerning practical employment of the EV-SSMP model to manage operations beyond the time horizon considered in this paper. Swap station managers could employ a *rolling horizon method* (Alden and Smith 1992) to generate and implement solutions to the infinite horizon, nonhomogeneous EV-SSMP MDP model. In one such implementation of the rolling horizon method, we fix the time horizon at $|T|$ periods (e.g., $|T| = 168$ to look ahead one week and capture day-of-the-week fluctuations), solve the corresponding $|T|$-period problem, implement the initial policy found during the next $T' < |T|$ periods (e.g., $T' = 24$ to use the day one decision rule to make charging/discharging decisions), roll forward T' periods (e.g., 24 periods), and repeat the method at the new current state. When repeating the process we incorporate pertinent seasonal and day-of-the-week data for the new, last day of the horizon. In this manner, the EV-SSMP can be used to manage ongoing operations. Indeed, in conjunction with a rolling horizon method, the EV-SSMP MDP model enables consideration of seasonal effects on prices and consumer behavior, and allows the prescription of policies that account for such information. One important caveat is that because the horizon used in the rolling horizon method is fixed and finite, the sequence of policies generated may not be optimal. Nonetheless, a series of rolling finite horizon solutions provides a reasonable approximate solution to an infinite horizon, nonhomogeneous EV-SSMP.

3. Theoretical Results and Methodology

In this section we first prove that an optimal nonincreasing monotone policy exists for the EV-SSMP when the demand is governed by a discrete nonincreasing distribution. Using this result, we describe exact solution methods and two heuristic benchmark policies.

3.1. Optimal Structural Properties

Determining if the optimal policy of a MDP contains structure, such as monotonicity, is significant due to the ease of implementation, appeal to decision makers, and the ability for faster computation time (Puterman 2005). When an optimal policy has a monotone structure, it can be solved with specialized and more efficient algorithms. As such, we wish to prove that our system contains a nonincreasing monotonic structure.

A policy π is said to be nonincreasing if for each $t = 1, \ldots, N-1$ and any pair of states $s_i, s_j \in S$ with $s_i < s_j$, it is true that $d_t^\pi(s_i) \geq d_t^\pi(s_j)$. We can demonstrate a nonincreasing monotone policy using a series of five properties regarding the reward function and the probability of moving to a higher state (Puterman 2005). Define

$$g_t(k|s_t, a_t) = \sum_{j \in \{S|j \geq k\}} p_t(j|s_t, a_t), \quad t = 1, \ldots, N-1 \tag{8}$$

as the probability of moving to state $j \geq k$ at decision epoch $t+1$ when action a_t is chosen in state s_t at decision epoch t. Let $A_{s_t} = A'$ for all $s_t \in S$, where $A' = \{\cup_{s_t \in S} A_{s_t}\}$ is the set of all possible actions independent of the state of the system. We note that a function, $f(x, y)$, is said to be subadditive (Puterman 2005) if for $x \geq \tilde{x}$ and $y \geq \tilde{y}$,

$$f(x, y) + f(\tilde{x}, \tilde{y}) \leq f(x, \tilde{y}) + f(\tilde{x}, y). \tag{9}$$

THEOREM 1. *There exists optimal decision rules $d_t^* : s_t \to A_{s_t}$ for the EV-SSMP that are nonincreasing in s_t for $t = 1, \ldots, N-1$ when demand D_t is governed by a nonincreasing discrete distribution.*

The claim is shown by demonstrating that the EV-SSMP exhibits the following 5 conditions (Puterman 2005).

1. $r_t(s_t, a_t)$ is nondecreasing in s_t for all $a_t \in A'$,
2. $g_t(k|s_t, a_t)$ is nondecreasing in s_t for all $k \in S$ and $a_t \in A'$,
3. $r_t(s_t, a_t)$ is a subadditive function on $S \times A'$,
4. $g_t(k|s_t, a_t)$ is a subadditive function on $S \times A'$ for all $k \in S$, and
5. $r_N(s_N)$ is nondecreasing in s_N.

Please see Appendix A for full details of the proof. In Theorem 1 it is proven that there exists optimal decision rules for the EV-SSMP that are nonincreasing in s_t when demand is governed by a nonincreasing discrete distribution. Consider two possibilities for the state (i.e., number of full batteries) at a swap station $s_t \geq \tilde{s}_t$. Through the proof of Theorem 1, we have shown that there exists optimal decision rules (i.e., optimal selected actions) such that the swap station will never charge less (or discharge more) batteries when in state state $\tilde{s}_t$ as compared to s_t. Utilizing this result, we outline exact solution methods and two benchmark solution methods.

3.2. Optimal Solution Methods

The objective in solving our Markov decision problem (MDP) is to determine a policy that maximizes the expected total reward criterion expressed in Equation (6). The set of states S is finite and the action set A_{s_t} is finite for each $s_t \in S$. Therefore there exists a deterministic Markov policy

that is optimal. We find an optimal policy for this finite horizon model by using the backward induction algorithm (Puterman 2005). This dynamic programming algorithm finds the optimal policy, or specifically, the optimal number of batteries to charge and discharge at each decision epoch, that maximizes the expected total reward. The backward induction algorithm finds sets $A^*_{s_t,t}$ that contain all actions in A_{s_t} that attain the maximum for the optimality equations (7). The algorithm also evaluates the policy and computes the expected total reward from each period to the end of the decision making horizon.

Our policy contains a nonincreasing monotonic structure when demand is governed by a discrete nonincreasing distribution, thus we also utilize the monotone backward induction algorithm (Puterman 2005) to find an optimal policy. The nonincreasing monotone backward induction algorithm modifies the original algorithm by redefining the action set at each iteration of s_t to be limited by the optimal decision rule of $s_t - 1$ for each $t \in T$. For example, if the optimal decision rule at $s_t = 10$ is to charge 20 batteries, then the action space for $s_t = 11$ will now be $A_{11} = \{\max(-11, -\Phi), \ldots, 0, \ldots, \min(20, \Phi)\}$ instead of $A_{11} = \{\max(-11, -\Phi), \ldots, 0, \ldots, \min(M-11, \Phi)\}$. The modifications to the algorithm will result in an optimal policy when demand is governed by a discrete nonincreasing distribution; note however, that there may be alternative optima that are not monotone.

When there are $|S|$ states, $|A'|$ actions in each state where $A' = \{\cup_{s_t \in S} A_{s_t}\}$, and N time periods, the backward induction algorithm requires $(N-1)|A'||S|^2$ multiplications to determine the optimal policy, which is a considerable improvement from complete enumeration of all possible solutions. Complete enumeration takes $(|A'|^{|S|})^{(N-1)}(N-1)|S|^2$ multiplications. In the worst case scenario, the monotone backward induction algorithm's computational effort equals that of the backward induction, however, when the policy is nonincreasing the action sets decrease in size with increasing s_t and reduce the number of actions that need to be evaluated (Puterman 2005).

3.3. Benchmark Policies

We consider two benchmark policies such that the swap station charges-up-to or discharges-down-to a set target level, ζ_t, at each decision epoch t. The design of these benchmark policies using a target level ensures the determined policy exhibits a monotone nonincreasing structure. The first benchmark policy is a stationary benchmark policy that picks a set target level ζ and sets $\zeta_t = \zeta$ for all time periods t. The second is a dynamic benchmark policy and utilizes a distinct ζ_t for each time period t. Utilizing each target level, the policy can be determined by calculating the action for each state and time period with a simple calculation. Thus, this policy can be easily implemented by the swap station manager.

If the state, s_t, is less than or equal to the target level, ζ_t, the swap station does not have as many fully charged batteries as desired, thus they will charge or do nothing. The most that can be charged at any point in time, denoted C, is given by

$$C = \min\{M - s_t, \Phi\}. \tag{10}$$

If s_t is greater than ζ_t the swap station has more fully charged batteries than desired, thus they will discharge. The most that can be discharged at any point in time (i.e., the most negative action), denoted D, is given by

$$D = \max\{-s_t, -\Phi\}. \tag{11}$$

The stationary benchmark decision rule $d_t^{\pi_s}(s_t)$ is then given by

$$d_t^{\pi_s}(s_t) = \begin{cases} \min\{\zeta - s_t, C\} & \text{if } s_t \leq \zeta, \\ \max\{\zeta - s_t, D\} & \text{if } s_t > \zeta, \end{cases} \tag{12}$$

and the dynamic benchmark decision rule $d_t^{\pi_d}(s_t)$ is given by

$$d_t^{\pi_d}(s_t) = \begin{cases} \min\{\zeta_t - s_t, C\} & \text{if } s_t \leq \zeta_t, \\ \max\{\zeta_t - s_t, D\} & \text{if } s_t > \zeta_t. \end{cases} \tag{13}$$

For the first benchmark policy π_s, we derive a stationary target level ζ, where ζ is calculated as a percentage of the number of batteries M using constant parameter $\mathscr{C}_s$. Equation (14) calculates ζ using a traditional rounding function. In the second benchmark policy π_d, we derive dynamic target levels ζ_t at each decision epoch as a rounded function of the number of batteries M, current and future charging costs, K_t and K_{t+1}, and the mean demand in the following time period, λ_{t+1}. The constant is a function of the charging cost in the current time period t and the next time period $t+1$. If the charging price will increase (i.e., $K_t \leq K_{t+1}$) it is desirable to charge as many batteries as possible in time period t. Contrarily, if the charging price is going to decrease (i.e., $K_t > K_{t+1}$) then it is desirable to charge more batteries in time period $t+1$. However, to ensure that demand is met even when it is desirable to charge in subsequent periods, we derive the target as a linear function of the mean demand in the next time period, λ_{t+1}, and the average number of swaps, γ, using constant parameter $\mathscr{C}_d$. Target levels are calculated using a traditional rounding function of the average number of swaps that will occur in the next time period (see Equation (15)).

$$\zeta = \lfloor \mathscr{C}_s M + 0.5 \rfloor \tag{14}$$

$$\zeta_t = \begin{cases} M & \text{if } K_t \leq K_{t+1} \\ \lfloor M \left(\frac{\mathscr{C}_d(\lambda_{t+1})}{\gamma} \right) + 0.5 \rfloor & \text{if } K_t > K_{t+1} \end{cases} \quad \forall t = 1, \ldots, N-1 \tag{15}$$

The performance of these benchmark policies will be sensitive to the selection of $\mathscr{C}_s$ and $\mathscr{C}_d$; thus, a computational study should be conducted for the data specific to each swap station. We explain our selection of these parameters and validate the benchmark policies in Section 4 as usable for real time decision making activities due to their speed of calculation and accuracy.

4. Computational Results

In this section, we computationally test the EV-SSMP on a variety of different scenarios. From the optimal policies, we deduce insights beneficial to a swap station manager. Further, we quantify the accuracy and speed of the two benchmark policies as compared to the optimal policy found via an optimal solution method.

The time horizon we examine is a full week in one hour increments. Thus, the time horizon is $N = (24)(7) + 1 = 169$ and the number of decision epochs is $N - 1 = 168$. The first decision is made on Monday at 0000, the second on Monday at 0100, until the last decision is made on Sunday at 2300. We utilize historical hourly charging cost data from 2013 in the Capital Region, New York, obtained from National Grid (2013). We use one week from each season in our analysis due to the varying climate and drastic variation in prices throughout the year. January 21-27 is used for Winter, April 15-21 for Spring, July 15-21 for Summer, and September 23-29 for Fall. Note that the sum of power prices over every hour of the week is at the maximum for January 21-27 and at a minimum for September 23-29 in 2013. The charging cost per kWh at each time t is multiplied by 60 to calculate the cost to charge one battery, K_t, which is consistent with the Tesla Model S 60 kWh battery option (Tesla motors 2014d). Charging can be completed in an hour with level 2 or 3 charging (Morrow et al. 2008) and is comparable to the Tesla supercharger option (Tesla motors 2015). The charging cost per battery per hour for the four weeks of interest is illustrated in Figure 2. For our computational tests, we set the discharge revenue, J_t, equal to a percentage of the charging cost, $J_t = \alpha K_t$, with α between 0.75 and 1.25. The α parameter enables examination of pricing mechanisms employed by the power company to incentivize a swap station to discharge at favorable points in times.

We consider a similar methodology to derive the distribution for swap demand at each hour as Nurre et al. (2014). The authors assume that the behaviors for arrivals at a swap station will mimic the currently observed behaviors at a gas station. Due to the fact that swap stations are being considered for adoption, we do not have actual swap station demand data. Thus, because one benefit of the swap station is convenience, we assume that the time when it is currently convenient to refill with gasoline will be similar to the convenient times for swapping. Nurre et al. (2014) calculate the percentage of people who will frequent a gas station for each hour of a day and day of a week based on historical data at Chevron gas stations (Nexant, Inc. et al. 2008). We utilize this

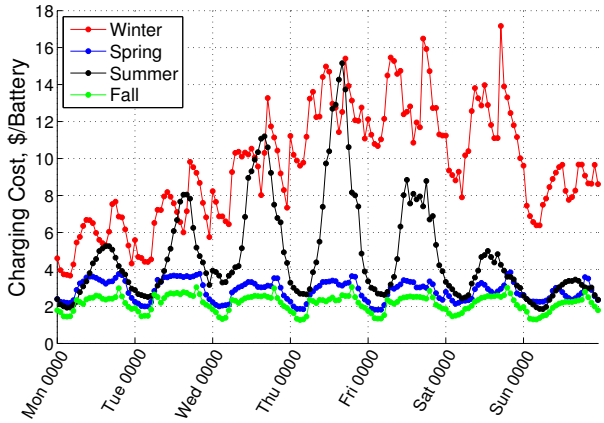


Figure 2 **Charging cost K_t per battery per hour in the Capital Region, NY.**

percentage to calculate the mean arrival rate of customers $\bar{X}_t$, for each decision epoch t. Specifically, we consider a time horizon with an average of γ swaps and set $\bar{X}_t$ equal to the product of γ and the percentage of swaps occurring at the station at time t from Nurre et al. (2014). As a simple example, one could assume that an EV makes one swap per time horizon and thus, γ represents the number of EVs in a location.

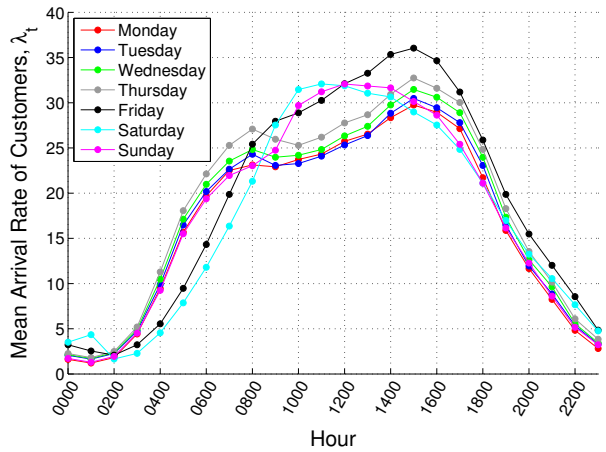


Figure 3 **Mean arrival rate of customers λ_t in a location with 3,000 swaps by hour and day of the week.**

We consider two distributions for modeling swap demand D_t: geometric and Poisson. When swap demand D_t follows a geometric distribution with parameter $\mathcal{P}_t$, we set $\mathcal{P}_t = \frac{1}{\bar{X}_t+1}$. When swap demand D_t follows a time dependent Poisson process with parameter λ_t, we set $\lambda_t = \bar{X}_t$. Note that the geometric distribution is a nonincreasing discrete distribution, therefore as was proven in Theorem 1, a monotonic nonincreasing policy is optimal. The mean arrival rate of customers $\lambda_t = \bar{X}_t$ for each hour of each day in a location with $\gamma = 3,000$ can be seen in Figure 3. We assume that the arrival rate is the same for each week of the year.

To computationally test the EV-SMMP, we conduct two designed experiments. The first designed experiment is conducted to gain insights when a wide range of inputs are considered. Specifically, we analyze the effects of non-controllable parameters, or external factors, such as demand, percentage earned from discharging batteries, and seasonal charging cost variations. The second designed experiment is conducted with targeted values based on the results of the first experiment. With this second experiment, we are able to determine how to appropriately set the controllable parameters, or internal factors, at a swap station such as swap price, number of batteries the swap station should have, and the charging capacity. With both, we utilize the expected total reward, percentage of met demand, and policies to infer policy insights.

4.1. Analysis of External Factors

We analyze the external factors, or those factors which will be outside of the direct control of the swap station, by conducting a Latin hypercube designed experiment with a wide range of inputs. With these results, we conduct a Monte Carlo simulation with three sample paths for observed demand to examine how uncertain demand impacts the operations throughout a typical week. Then, we address the incentives necessary to be offered to the swap station by the power company to encourage favorable discharging while simultaneously meeting demand. Finally, we analyze how the drastic variation of seasonal charging costs affects the expected total reward and the percentage of demand that is met.

4.1.1. First Latin Hypercube Designed Experiment. We perform a 50-scenario Latin hypercube designed experiment, which is a widely used design for deterministic computer simulation models (Montgomery 2008). This space filling design spreads the design points nearly uniformly to better characterize the response surface in the region of experimentation. For this designed experiment, we find the expected total reward using the monotone dynamic programming algorithm when demand is geometric. When demand follows a time dependent Poisson process we calculate two policies with corresponding expected total rewards: the optimal policy is found using the backward induction algorithm, and a heuristic policy is found using the monotone backward induction algorithm. We note, that the monotone policy is not always optimal when demand follows a time dependent Poisson process, however empirically we have observed it to be optimal or near-optimal in the majority of cases. Because we look at four separate weeks for charging cost data, we consider K_t a categorical factor with four levels representing the four weeks extracted from the year. We conduct the 50-scenario design for each of the four seasons and each of the two demand distributions, resulting in a total of 400 scenarios. Factors examined in the design include the total number of batteries M, the charging capacity Φ, the average number of swaps in the local area γ, the revenue per battery swap ρ, and the percentage α of revenue earned from

discharging with respect to the charging cost. Using JMP11Pro software, we generate a 50-scenario design with various levels of each factor ranging between two values. The high and low levels used for this experiment are displayed in Table 1. The charging costs for the four weeks of interest, $K_t^W, K_t^{Sp}, K_t^{Su}$, and K_t^F, are representative of Winter, Spring, Summer, and Fall, respectively. We set the low value for the swap revenue ρ to less than the minimum charging cost over the four weeks and the high value for ρ to greater than the maximum charging cost.

Table 1 Factor levels for first Latin hypercube designed experiment.

Factor		Low	High
Total Number of Batteries	M	50	200
Charging Capacity	Φ	$\lfloor 0.25M \rfloor$	M
Swap Revenue ($)	ρ	1	20
Percent Discharge Revenue ($\%K_t$)	α	0.75	1.25
Avg. Swaps in the Local Area	γ	1,000	6,000

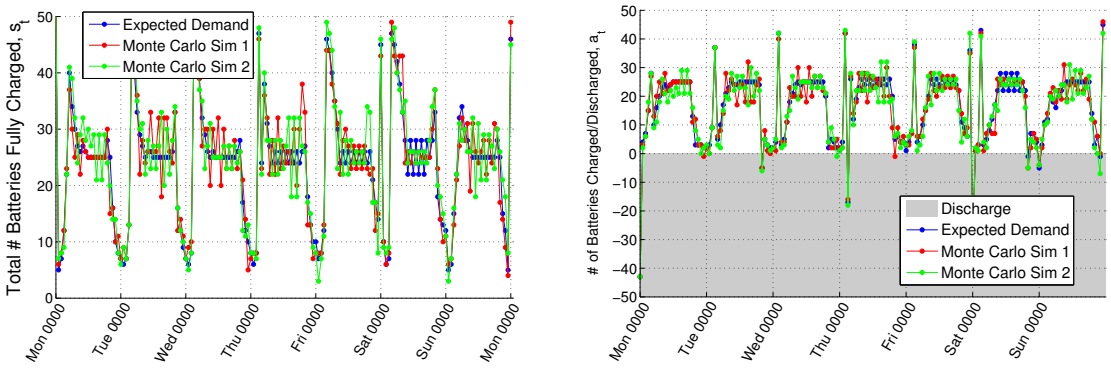
When considering the time dependent Poisson process for demand, the monotone policy was optimal in 128 out of 200 scenarios. Of the 72 scenarios that were not optimal, 48 of them were within 1% of optimality. Of the 24 scenarios not within 1% of optimality, the average percentage gap was 8.67%. Therefore, while the monotone policy is not always optimal when demand does not follow a nonincreasing distribution, we empirically observe that it provides a good approximation. Further, we observe very similar optimal policies when using Poisson and geometric demand. Discharging is more prevalent when demand follows a Poisson process, however, discharging does occur when demand is governed by a geometric distribution. Due to the similarities seen, the results presented herein apply to both distributions unless otherwise stated.

The results from this experiment, when demand follows a time dependent Poisson process and the monotone backward induction algorithm is used, indicate that all factors have a statistically significant effect on the expected total reward at the 95% confidence level. As expected, the swap revenue ρ has the greatest impact on the expected total reward. Thus, the most effective way to increase profit would be to increase the swap price, however this is based on the assumption that demand for swaps is independent of the swap price, which is unrealistic. Future work should consider the sensitivity of customer demand to the price for swapping, as utilizing a charging station can occur instead of swapping. We also note that although the charging capacity Φ is statistically significant, the effect on the profit is small compared to the other factors.

4.1.2. Effects of Uncertain Demand on Swap Station Operations. In this set of analyses, we seek to characterize the day-to-day operations at the swap station when faced with uncertain demand. We illustrate the state of the system, or the number of fully charged batteries the swap station has on hand, when operating using the monotone structured policy throughout a typical

week for a swap station with $M = 50$, $\Phi = M$, $\rho = 15$, $\gamma = 3,000$, $\alpha = 1$, and $K_t = K_t^{Sp}$. We note that in this scenario the nonincreasing monotone policy results in an optimal policy. We generate three sample paths for observed demand at the swap station. In the first sample path, we assume that the demand observed at the swap station is exactly the mean arrival $\lceil \lambda_t \rceil$ when demand follows a Poisson process. We then use Monte Carlo simulation to generate two sample paths for observed demand at each decision epoch based off the time dependent Poisson process and known mean arrival rate. We calculate the state at the next decision epoch $t+1$ using the monotone policy decision rule $d_t^*(s_t)$ for the current state s_t and time t and the observed demand, denoted $\hat{X}_t$, by way of

$$s_{t+1} = s_t + d_t^*(s_t) - \min\left\{\hat{X}_t, s_t - |\min\{d_t^*(s_t), 0\}|\right\}. \tag{16}$$

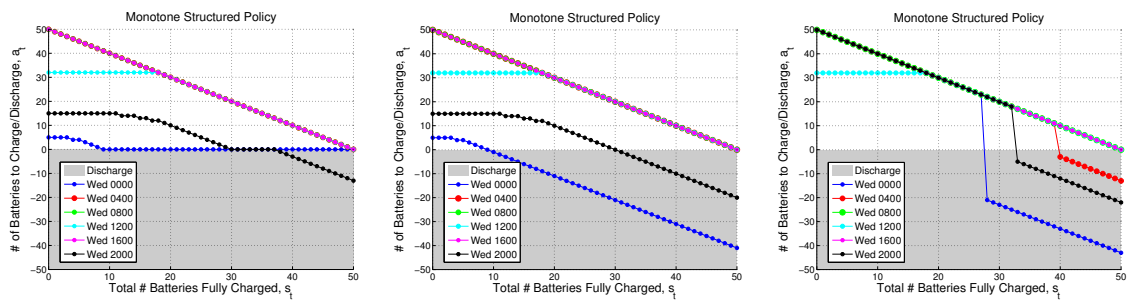


(a) State versus time. (b) Action versus time.

Figure 4 State and action over a week time period for three simulated observed demands.

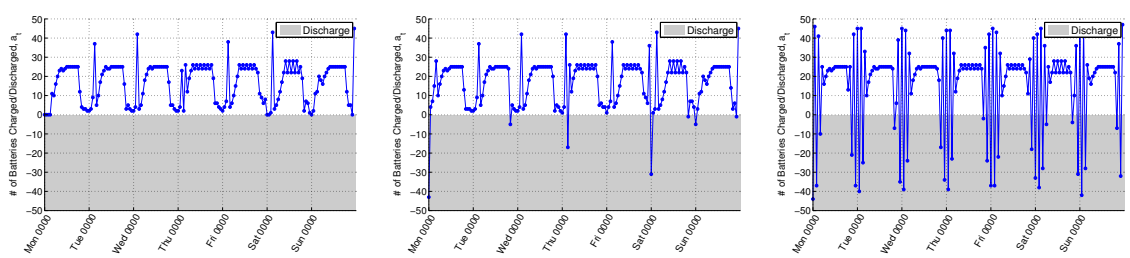
Assuming the swap station starts with all full batteries, we examine an entire week. The state of the system at each decision epoch and the corresponding action are shown in Figure 4. From this figure, we first note that our assumption that the swap station starts with all full batteries at the beginning of a time horizon is not a simplifying assumption as the number of full batteries naturally increases at the start of each day. We also note that the state of full batteries stabilizes at around 25 batteries in the mid-late afternoon each day. Moreover, the state and action taken is relatively consistent for each of the three observed sample paths of demand. This is a desirable result, leading to the **policy insight** that the action taken in relation to the state balances. In the observed scenario, when the sample path for observed demand is the mean arrival $\lceil \lambda_t \rceil$, 91.25% of demand is met when this policy is implemented. Our second designed experiment addresses the relationship between the internal controllable parameters to ensure an acceptable level of demand is met.

4.1.3. Percentage Earned for Discharging Insights. We proceed with our analysis by seeking to determine what incentive (α) the power company will provide the swap station to encourage favorable discharging behavior. We examine the monotone structured policies for different scenarios when demand follows a Poisson process. In Figure 5 we illustrate the policies for a scenario with $M = 50$, $\Phi = M$, $\rho = 15$, $\gamma = 3,000$, and $K_t = K_t^{Sp}$ differentiated by three values for α. We note that in this scenario the nonincreasing monotone policy results in an optimal policy for each of the α values. For a typical Wednesday, Figures 5a, 5b, and 5c show the policies in 4 hour increments. For $\alpha = 0.75$, we observe that the monotone policy rarely indicates to discharge. The optimal action rarely drops below zero (the grayed area of Figure 5a). For $\alpha = 1$ discharging does occur when the number of full batteries at the swap station is above some inventory threshold; this threshold varies for each time period. For $\alpha = 1.25$, the optimal policy alternates between charging and discharging when the swap station has about 28 or more fully charged batteries, as can be seen in Figure 5c.



(a) $\alpha = 0.75$ (b) $\alpha = 1$ (c) $\alpha = 1.25$

Figure 5 Monotone policy by percentage of the charge cost earned for discharging, α.



(a) $\alpha = 0.75$ (b) $\alpha = 1$ (c) $\alpha = 1.25$

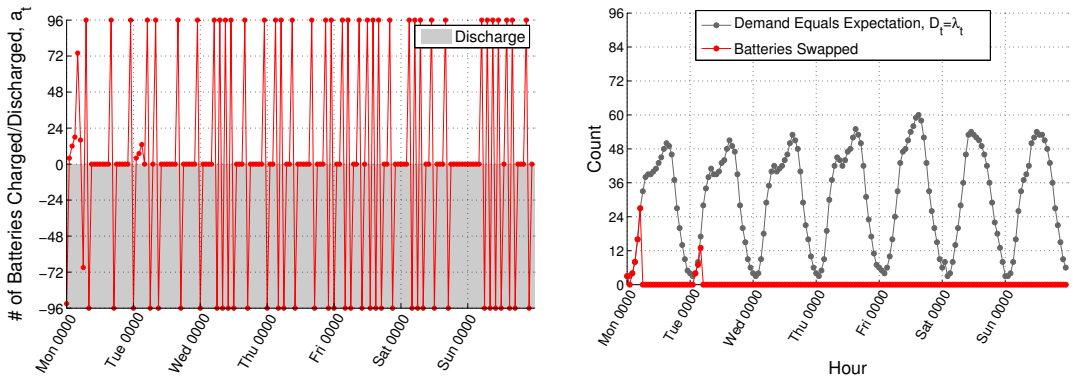
Figure 6 Actions taken throughout a typical week by percentage of the charge cost earned for discharging, α.

We further examine the actions that will be taken throughout a typical week based on the number of fully charged batteries available in each time period. To do this, we simulate a typical week assuming that the demand observed at the swap station is exactly the mean arrival $\lceil \lambda_t \rceil$. The

actions taken throughout the week differentiated by three values for α can be seen in Figure 6. When $\alpha = 0.75$, we observe in Figure 6a that discharging does not occur for the entire week even though the policy in Figure 5a does indicate discharging. Further, when $\alpha = 1.25$, the effect of alternating between charging and discharging in the policy in Figure 5c can be seen by the oscillation in Figure 6c. From this observation, we examine all policies from the designed experiment with $\alpha < 1$ and $\alpha > 1$ to see if the policies mimic those when $\alpha = 0.75$ and $\alpha = 1.25$, respectively. In general, we notice that when $\alpha < 1$ discharging is much less prevalent, unless there are significantly more batteries than needed to meet demand or ρ is set too low compared to the charging costs, resulting in the situation where discharging earns more revenue than meeting swap demand. The swap station must consider the cost to have significantly more batteries than needed to meet demand to aid the power grid by discharging at a discounted rate. When $\alpha > 1$ the oscillating trend of charging and discharging is present. The negative behavior of oscillating between charging and discharging in consecutive time periods could lead to further variability in the power grid. Thus, from the power companies perspective, when $\alpha = 1$ the swap station exhibits a good balance between charging and discharging. From this analysis concerning α, we have deduced the **policy insight** that to maintain the dual purpose of the swap station of meeting swap demand and still exhibiting some favorable V2G discharging behavior that the money earned from discharging should exactly equal the charging cost, which is attained when $\alpha = 1$. Thus, in our further analysis we focus on the scenarios when $\alpha = 1$.

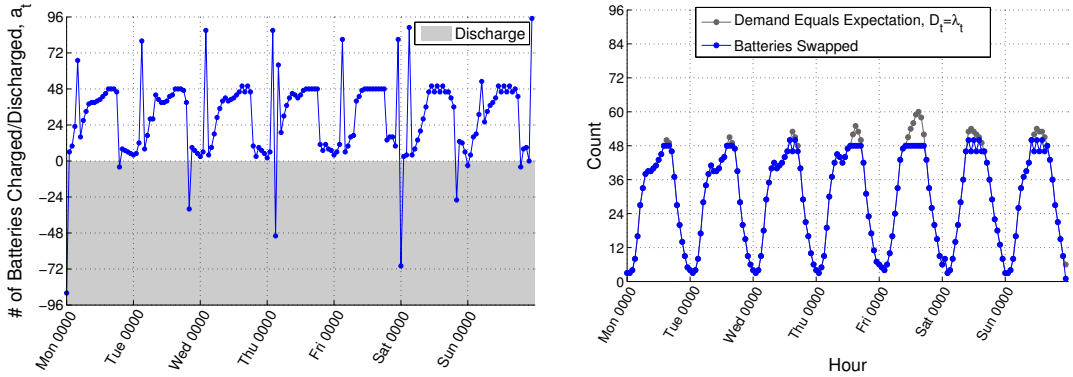
4.1.4. Seasonal Variations. Next, we seek to characterize the day-to-day operations at the swap station when faced with highly variable charging cost data by season. For each of the four seasons, we examine the monotone structured policy when demand follows a Poisson process for the scenario with $M = 96$, $\Phi = M$, $\rho = 4.88$, $\gamma = 4,980$, and $\alpha = 1$. We note that for each season in this scenario, the nonincreasing monotone policy results in an optimal policy. For each season, we simulate a typical week assuming that the demand observed at the swap station is exactly the mean arrival $\lceil \lambda_t \rceil$. Note, the average charging costs are \$9.58, \$2.86, \$4.80, and \$2.17 for Winter, Spring, Summer, and Fall, respectively and the swap price in this scenario is \$4.88.

Figure 7 illustrates the actions taken and the number of batteries swapped compared to demand for Winter. In this case, ρ is too low relative to the charging costs for the swap station to meet demand over the opportunity cost to discharge to earn revenue. Thus, the actions taken throughout the day alternate between charging and discharging (see Figure 7a) and only 1.62% of demand is met (see Figure 7b). Figure 8 illustrates the actions taken and the number of batteries swapped compared to demand for Spring. In this case, ρ is appropriately set relative to the charging costs. The swap station is motivated to meet demand over discharging to earn revenue. The actions taken



(a) Action versus time. (b) Demand compared to batteries swapped.

Figure 7 Action and met demand over a week for simulated observed demand in Winter.

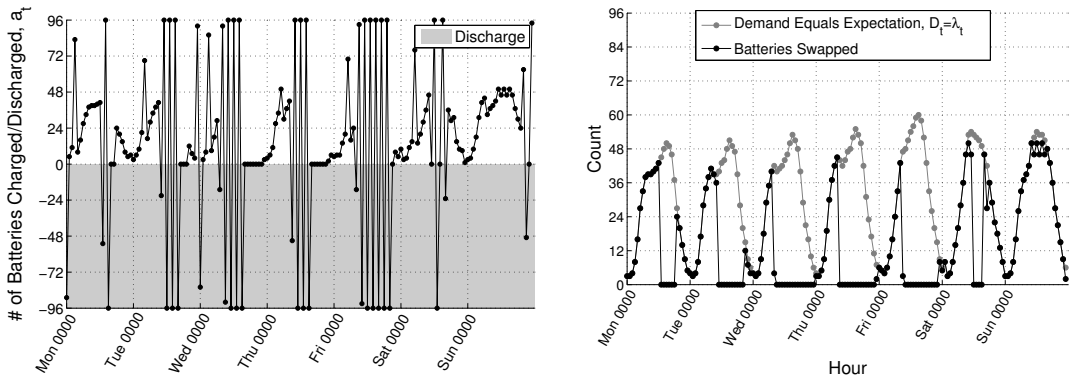


(a) Action versus time. (b) Demand compared to batteries swapped.

Figure 8 Action and met demand over a week for simulated observed demand in Spring.

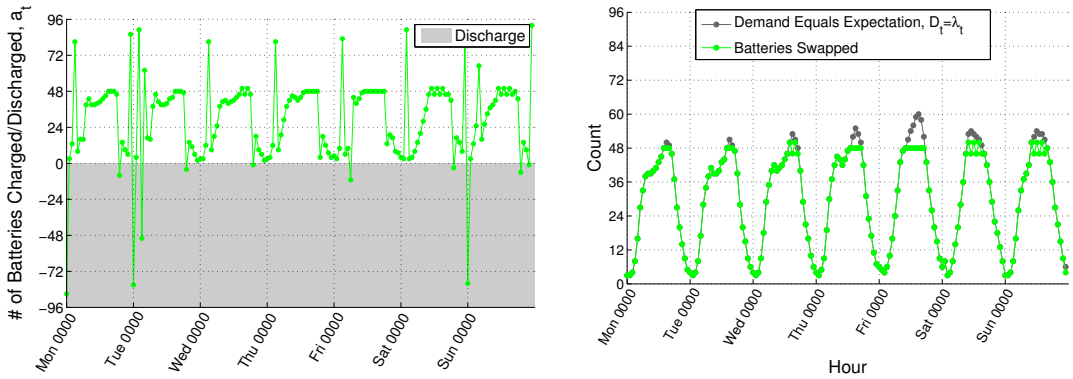
throughout the day are more stabilized (see Figure 8a) than in Winter and 97.18% of demand is met (see Figure 8b). Figure 9 illustrates the actions taken and the number of batteries swapped compared to demand for Summer. In this case, ρ is set higher than the average charging cost, however the highly variable prices in Summer cause alternating between charging and discharging (see Figure 9a) and 46.34% of demand is met (see Figure 9b). While this is an improvement over Winter, these results indicate a higher ρ is necessary in Summer or the ρ value should vary by time of day. Figure 10 illustrates the actions taken and the number of batteries swapped compared to demand for Fall. We see similar results as in Spring due to very similar charging costs. The swap station is motivated to meet demand over discharging to earn revenue (see Figure 10a) and 97.24% of demand is met (see Figure 10b).

For each season, we notice that when less demand is being met, the expected total reward is less. The expected total rewards for this scenario are \$6,042.66, \$9,754.82, \$6,478.43, and \$13,128.20 for Winter, Spring, Summer, and Fall, respectively. From these results, we deduce the **policy insight** that it is necessary to set the swap price ρ appropriately based on the seasonal charging costs to



(a) Action versus time. (b) Demand compared to batteries swapped.

Figure 9 Action and met demand over a week for simulated observed demand in Summer.



(a) Action versus time. (b) Demand compared to batteries swapped.

Figure 10 Action and met demand over a week for simulated observed demand in Fall.

meet demand over the opportunity cost to discharge. We address this in our second experiment by using different ranges for ρ in each season.

4.2. Analysis of Internal Factors

We analyze the internal factors by conducting a Latin hypercube designed experiment with more targeted values based on insights drawn from the first experiment. These insights include setting the discharge revenue exactly equal to the charging cost for desirable discharging and satisfaction of demand and setting ρ appropriately with respect to the seasonal charging costs. The goal of the second experiment is to draw insights on the appropriate levels of the controllable parameters at the swap station, or synonymously the internal factors.

4.2.1. Second Latin Hypercube Designed Experiment. For the second designed experiment, we seek to determine the appropriate levels for internal factors in relation to the non-controllable external factors. Specifically, we seek to gain insights into acceptable levels for the number of batteries M, charging capacity Φ, and swap price ρ. With this, we broaden our focus

to include, in addition to the expected total reward, the expected amount of demand that is met. We perform a second Latin hypercube designed experiment with 40 scenarios, which is generated using JMP11Pro software. The high and low levels used for this experiment are shown in Table 2. Different high and low values for ρ are used in each season. The lower bound on ρ is found by rounding the average charging cost to the nearest dollar in each season and the upper bound is set to approximately 1.5 times the maximum charging cost.

Table 2 **Factor levels used for the second Latin hypercube designed experiment.**

Factor		Low	High
Total Number of Batteries	M	50	200
Charging Capacity	Φ	$\lfloor 0.25M \rfloor$	M
Winter Swap Revenue ($)	ρ_W	10	25
Spring Swap Revenue ($)	ρ_{Sp}	3	10
Summer Swap Revenue ($)	ρ_{Su}	5	20
Fall Swap Revenue ($)	ρ_F	2	8
Avg. Swaps in the Local Area	γ	$1,000$	$6,000$

We execute this experiment for two cases: when demand follows a geometric distribution and when demand follows a Poisson process. Both a monotone policy and an optimal policy are found when demand follows a Poisson process. Because Poisson is not a monotonic nonincreasing distribution, the monotone policy is not guaranteed to be optimal. However, when demand follows a time dependent Poisson process, we found that for the 160 scenarios in this experiment, the monotone policy is optimal in all but 18 scenarios. Of the 18 scenarios that are not optimal, the maximum percentage gap is 1.84%. This leads us to the **policy insight** that a monotone policy is a good approximation for the optimal policy even when demand is not governed by a nonincreasing distribution.

The expected total reward found for all scenarios are very similar when demand follows a Poisson process and when demand follows the geometric distribution. When comparing the expected total reward found for the same scenarios, solved with demand following a geometric distribution and a Poisson distribution, a total of 153 out of 160 scenarios were within 10% of each other. Due to these similarities, we focus on presenting the monotone policies found for scenarios when demand follows a Poisson process.

When examining the results from this second designed experiment, we find that M, ρ, γ, and seasonal charging cost, K_t have a significant effect on the expected total reward at the 95% confidence level. When analyzing the expected percent of demand that is met when the monotone structured policy is implemented, we conclude that M, γ, and ρ are significant. We proceed with our analysis by further examining the policies found under different scenarios with the goal of gaining insight into the acceptable levels for the controllable internal factors at a swap station.

4.2.2. Charging Capacity Insights. By examining the utilization of the charging capacity, percentage of demand met, and expected total reward under different circumstances, we seek to gain insight into the number of physical plug-ins (i.e., charging capacity) the swap station should install. We first note that the charging capacity Φ was found to have a slight effect in the first designed experiment. However, when we focus on $\alpha = 1$ and season specific ρ we find that charging capacity does not have a significant effect on the expected total reward, which leads to the **policy insight** that the swap station has flexibility when designing the charging infrastructure at the swap station. For all seasons when the charging capacity is higher, the swap station uses the full charging capacity available, which does not ultimately increase profit. Whether or not the swap station is utilizing a small charging capacity or a large charging capacity, the percentage of demand that the swap station is able to meet remains unaffected. One driving factor in the satisfaction of demand is the number of batteries at the swap station, which we will discuss in Section 4.2.4. Further, long-term analysis needs to be conducted to better inform the charging capacity that should be installed at each swap station. This analysis should include forecasts for the change in adoption of EVs over time in comparison to the price for expansions or future installations of charging capacity.

4.2.3. Effect of Swap Price on Meeting Demand. This next set of analyses seeks to determine how the price charged per battery swap impacts the percentage of met demand. As was demonstrated in Section 4.1.4, the swap price charged is in direct competition with the discharging revenue that could be earned. Thus, we examined swap prices in different ranges based on season. The seasonal charging cost K_t was found to be not significant with respect to percentage of demand met. This indicates the season specific swap prices selected mitigated the issues seen in Section 4.1.4 when discharging occurred instead of meeting swap demand when the discharging revenues (i.e., charging costs) were high, such as in Winter and Summer. Further, increasing the swap price ρ motivates the satisfaction of more demand over discharging for all seasons. We deduce the **policy insight** that in each season to meet over 95% of demand, a swap price of approximately \$20, \$6, \$14, and \$5 is desirable for Winter, Spring, Summer, and Fall, respectively. These prices are in line with the maximum charging cost for Winter and Summer but are slightly higher than the maximum for Spring and Fall. This difference by season is a direct result of the range of charge cost for each season. For both Winter and Summer, the range of charge cost is large; thus, when the swap price is set to the maximum charge cost and the charge cost is low, an increased amount of revenue can be earned by balancing out the times when swap price equals the maximum charge cost. In contrast, for Spring and Fall the ranges for charge cost are small leading to the result that the swap price needs to be slightly higher than the maximum charge cost.

4.2.4. Number of Batteries Insights. Next, we focus on determining a threshold level of batteries that are needed at the swap station to meet demand. Because ρ has a strong impact on the percentage of demand that is met at a swap station, in this analysis we deliberately set ρ to deduce the impact of M on the percentage of met demand. Specifically, we set ρ to be \$20, \$6, \$14, and \$5 for Winter, Spring, Summer, and Fall, respectively, which are the values found as a result of the analysis presented in Section 4.2.3. Using the parameters from the second designed experiment at the constant ρ value in each season, we determine that for this data set the swap station should have approximately $M = 3\%\gamma$ batteries or more at the swap station to meet at least 95% of demand. This leads to the **policy insight** that the number of batteries M must be in line with the average number of swaps in the local area γ to be able to satisfy demand.

4.3. Validation of Benchmark Policies

In this subsection, we examine the benchmark policies to assess their accuracy and speed. For all scenarios in the second Latin hypercube experiment, we test the stationary benchmark policy π_s with $\mathscr{C}_s = 0.8$ (see Equation (17)) and the dynamic benchmark policy π_d with $\mathscr{C}_d = 100$ using Equation (18). In the stationary policy, the constant $\mathscr{C}_s$ was chosen with the goal that the swap station should have 80% of their batteries fully charged at any point in time. As we saw in the analysis presented in Section 4.1.2, the number of full batteries levels out at around 50% of full batteries. However, earlier times of each day have more full batteries. Thus, 0.8 was selected so that the swap station can ensure a high level of met demand by aiming to have a high number of batteries available for swapping. We incrementally increased from 0.5 upward and found 0.8 to perform best for the stationary benchmark policy.

However, as we have seen in analysis on the day-to-day operations of the swap station, the state of the system is highly dependent on time. Thus, for the dynamic benchmark policy, the constant $\mathscr{C}_d$ was chosen so that the swap station would have a sufficient number of batteries to meet the mean demand in the following period without charging more than needed, if the cost to charge is going to decrease. If the cost to charge is going to increase in time $t+1$, the swap station should charge as much as possible in time t at a cheaper price regardless of demand. We found that setting $\mathscr{C}_d = 100$ accomplished this goal. With $\mathscr{C}_d = 100$ in Equation (18), the resulting target is exactly the percentage of mean battery swaps occurring in the next time period in relationship to the mean swaps over the time horizon times 100%. As was done in the stationary benchmark policy, we tested for values higher and lower than $\mathscr{C}_d = 100$ and found this selection to empirically perform best. We note, however, that analysis should be conducted with different data sets, as the performance of the benchmark policies is sensitive to the selection of $\mathscr{C}_s$ and $\mathscr{C}_d$.

$$\zeta = \lfloor 0.8M + 0.5 \rfloor \tag{17}$$

$$\zeta_t = \begin{cases} M & \text{if } K_t \leq K_{t+1} \\ \lfloor M\left(\frac{100(\lambda_{t+1})}{\gamma}\right) + 0.5 \rfloor & \text{if } K_t > K_{t+1} \end{cases} \quad t = 1, \ldots, N-1 \tag{18}$$

For all tests, we compare the computation time, optimal expected total reward, and expected percentage of met demand to an optimal policy found via the backward induction algorithm (BI). An optimality gap is calculated using the optimal expected total reward $\upsilon_N^*(s_1)$ and found expected total reward $\upsilon_N^\pi(s_1)$ for policy π using Equation (19), where an optimality gap of 0.00% indicates an optimal solution has been found.

$$\text{Optimality Gap} = \frac{\upsilon_N^*(s_1) - \upsilon_N^\pi(s_1)}{\upsilon_N^*(s_1)} \tag{19}$$

The expected percentage of demand met is compared by calculating a demand gap equal to the subtraction of the value found in the benchmark policy from the value found in the optimal policy. With this calculation, a positive number indicates that the optimal policy is meeting more demand, whereas a negative number indicates the benchmark policy is meeting more demand.

For the stationary benchmark policy, the average optimality gap is 20.33% when we exclude scenarios with unrealistically low swap prices. The average demand gap is −1.28%, meaning on average the stationary benchmark policy meets 1.28% more demand than the optimal policy. At best, the π_s policy meets 32.63% more demand and at worst π_s meets 12.26% less demand than the optimal policy. For the dynamic benchmark policy, the average optimality gap is 17.76% when we don't include scenarios with unrealistically low swap prices. The average demand gap is −2.35%, meaning on average the dynamic benchmark policy meets 2.35% more demand than the optimal policy. At best, the policy meets 3.75% more demand and at worst π_d meets 8.64% less demand than the optimal policy. The full results of the benchmark policies for Winter, Spring, Summer, and Fall are presented in Appendix B.

These results indicate that the dynamic benchmark policy outperforms the stationary benchmark policy due to decreased optimality gaps and demand gaps. Specifically, we found the π_d policy to have decreased optimality gaps and much smaller demand gaps than the π_s policy over all scenarios. Since many of the parameters are unknown, the result of these benchmark policies perform well over a wide range of possibilities. For future work we recommend additional analysis of potential benchmark policies when more of the parameters are known. The swap station must also consider the cost of the capability to solve the model to optimality versus the loss in profit from implementing a benchmark policy.

Optimal solution methods require the use of probability transition matrices and reward vectors to calculate the optimal policy, while benchmark policies only require probability matrices and

reward vectors to *evaluate* the policy, not calculate the implementable action. The average times for creating the probability matrices and reward vectors were 1921.94 and 18.58 seconds, respectively. The average computation time for the backward induction algorithm was 27.86 seconds. The average computation time for the benchmark policies was 4.03 seconds. Computations were done using MATLAB R2014a software on a 2.4 GHz Intel Core i5 processor laptop with 4GB 1600 MHz DDR3 of memory.

We deduce the **policy insight** that the dynamic benchmark policy could be a viable option for implementation at a swap station. This benchmark policy allows for an easy calculation and subsequent implementation at the swap station of the number of batteries to charge and discharge over time based off a target level for each hour of a week. Therefore, all that is needed is 168 target values; one value for each hour of the week. In contrast, implementation of the optimal policy would require a very large look up table by state and time. For a modest amount of 50 batteries, there is a corresponding look up table of $168(M) = 168(50) = 8,400$ values. Further, we hypothesize that the smaller amount of numbers and easily described benchmark policies could increase the user trust in the output solution.

4.4. Summary of Policy Insights

We summarize the results and analysis of these computational tests in the following policy insights for an EV swap station manager and the power grid.

1. The state and action taken throughout a week is consistent. Thus, we conclude the action taken in relation to the state balances, enabling the desirable property of consistent status of the swap station (i.e., number of full batteries) by day even when faced with uncertainty.
2. When the incentive to discharge is too high, the negative behavior of oscillating between charging and discharging in consecutive time periods occurs at the swap station, thereby leading to further variability in the power grid. When the incentive is too low and ρ is set appropriately, discharging rarely occurs. We conclude that $\alpha = 1$ (the revenue earned from discharging is exactly the cost for charging) results in a good balance of some discharging but limited oscillating behavior, thereby enabling the dual purpose of the swap station.
3. To ensure that the swap station is meeting demand and not solely focused on discharging to earn revenue, the swap price must be set appropriately with respect to the seasonal charging costs. For at least 95% of demand to be met on average, the swap station should set the swap price equal to the maximum charging cost for Winter and Summer, and slightly higher than the maximum for Spring and Fall.
4. When the distribution for demand at the swap station does not have a nonincreasing structure, the monotone policy provides a good approximate solution. This allows for ease of implementation of the structured policy and provides a basis for the development of future benchmark policies.

5. An EV swap station has flexibility when designing the physical swap station charging capacity. When the revenue earned from discharging is exactly the cost for charging the expected total reward and percentage of demand that is met is unaffected by charging capacity. A greater charging capacity will be used if available but does not ultimately have a significant effect on profit or meeting demand as indicated by our region of experimentation. Future work should examine charging capacity values not correlated with the number of batteries.
6. It is integral to have the number of batteries at your swap station M in line with the average number of swaps or EVs in the local area γ for meeting demand, maximizing expected total reward, and allowing for discharging back to the power grid using V2G. We determine that for this data set the swap station should have approximately $M = 3\%\gamma$ batteries or more at the swap station to ensure 95% of demand is met on average.
7. The dynamic benchmark policy that calculates a target level for each time period in a time horizon was superior to a stationary benchmark policy. The action for the dynamic benchmark policy is to charge-up-to or discharge-down-to this time dependent target level based on the number of full batteries on hand.
8. For all scenarios considering different number of batteries, charging capacity, swap revenue, charging cost by week, incentive to discharge, and average number of swaps per week in a local area, the swap station was always able to remain profitable with our model. Certain combinations of these factors led to greater profitability, but this result indicates that in all circumstances considered, a swap station is a viable, profitable option for EVs.

5. Conclusions

We have considered the problem of managing the operations at an electric vehicle swap station when demand for swaps is uncertain. In this context, we seek to determine how many batteries the swap station should charge and discharge (utilizing V2G) over time. We model this problem using a Markov decision process when demand follows a discrete distribution. Further, we proved that there exists an optimal nonincreasing monotone policy when demand follows a discrete nonincreasing distribution. Therefore, both the backward induction and monotone backward induction algorithms can be utilized to find the optimal policy. We created two easy-to-implement benchmark policies and empirically compared their performance to an optimal policy. Two designed experiments were performed, from which we deduced many insights including: (i) the dynamic benchmark policy is best; (ii) the swap price must be in line with the seasonal charging costs; (iii) the number of batteries is an integral parameter for meeting demand; and (iv) α needs to be appropriately set by the power company to encourage discharging and not oscillating behavior. Future work should consider how the swap price impacts the demand for swaps in

comparison to using at home charging or a charging station. Analysis should also be conducted to quantify the cost of not meeting swap demand when insufficient inventory is available or discharging using V2G is more profitable than swapping. Further, uncertainties regarding power prices, power load, and other renewables should be incorporated into the state space of the MDP to fully capture the load balancing potential of an EV swap station. Additionally, models should be created that capture the true state of the batteries exchanged and the corresponding time needed to fully charge or discharge these batteries.

Disclaimer. The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

Acknowledgments

The authors thank the Editor, Associate Editor, and two anonymous referees for their constructive comments and suggestions that have greatly helped improve the substance and presentation of this paper.

References

Abreu, P. 2010. *The world's only electric sports car: 2010 Tesla Roadster*. Motor Authority. Last accessed on May 20, 2015 at `http://www.motorauthority.com/news/1044161_the-worlds-only-electric-sports-car-2010-tesla-roadster`.

Alden, JM, RL Smith. 1992. Rolling horizon procedures in nonhomogeneous Markov decision processes. *Operations Research* **40**(2) S183–S194.

Bingliang, Z, S Yutian, L Bingqiang, L Jianxiang. 2012. A modeling method for the power demand of electric vehicles based on Monte Carlo simulation. *2012 Asia-Pacific Power and Energy Engineering Conference*. Shanghai, China, 1–5.

Clement-Nyns, K., E. Haesen, J. Driesen. 2010. The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. *IEEE Transactions on Power Systems* **25**(1) 371–380.

ElHafsi, M. 2009. Optimal integrated production and inventory control of an assemble-to-order system with multiple non-unitary demand classes. *European Journal of Operational Research* **194**(1) 127–142.

Eyer, J., G. Corey. 2010. Energy storage for the electricity grid: Benefits and market potential assessment guide. Tech. rep., Sandia National Laboratories.

Fowler, S. 2014. Tesla model 3 to challenge BMW 3 series - World exclusive. Last accessed on May 20, 2015 at `http://www.autoexpress.co.uk/tesla/87867/tesla-model-3-to-challenge-bmw-3-series-world-exclusive`.

Ghate, A, RL Smith. 2013. A linear programming approach to nonstationary infinite-horizon markov decision processes. *Operations Research* **61**(2) 413–425.

Giannoccaro, I, P Pontrandolfo. 2002. Inventory management in supply chains: a reinforcement learning approach. *International Journal of Production Economics* **78**(2) 153–161.

Göransson, L, S Karlsson, F Johnsson. 2010. Integration of plug-in hybrid electric vehicles in a regional wind-thermal power system. *Energy Policy* **38**(10) 5482–5492.

Lewis, BM. 2005. Inventory control with risk of major supply chain disruptions. Ph.D. thesis, Georgia Institute of Technology.

Mak, H-Y, Y Rong, Z-JM Shen. 2013. Infrastructure planning for electric vehicles with battery swapping. *Management Science* **59**(7) 1557–1575.

Montgomery, DC. 2008. *Design and analysis of experiments*. 8th ed. John Wiley & Sons, New Jersey.

Morrow, K, D Karner, J Francfort. 2008. Plug-in hybrid electric vehicle charging infrastructure review. Tech. rep., U.S. Department of Energy, Idaho National Laboratory.

National Grid. 2013. Hourly electric supply charges. Last accessed on October 8, 2014 at `http://www.nationalgridus.com/niagaramohawk/business/rates/5_hour_charge.asp`.

Nexant, Inc., Air Liquide, Argonne National Laboratory, Chevron Technology Venture, Gas Technology Institute, National Renewable Energy Laboratory, Pacific Northwest Laboratory, TIAX LLC. 2008. H2A hydrogen delivery infrastructure analysis models and conventional pathway options analysis results. Last accessed on May 20, 2015 at `http://www1.eere.energy.gov/hydrogenandfuelcells/pdfs/nexant_h2a.pdf`.

Nurre, SG, R Bent, F Pan, TC Sharkey. 2014. Managing operations of plug-in hybrid electric vehicle (PHEV) exchange stations for use with a smart grid. *Energy Policy* **67** 364–377.

Pan, F, R Bent, A Berscheid, D Izraelevitz. 2010. Locating PHEV exchange stations in V2G. *2010 First IEEE International Conference on Smart Grid Communications*. Gaithersburg, MD, 173–178.

Pearson, D, ST Stub. 2013. *Better Place's failure is blow to Renault*. The Wall Street Journal. Last accessed on November 20, 2014 at `http://online.wsj.com/articles/SB10001424127887323855804578507263247107312`.

Peng, M, L Liu, C Jiang. 2012. A review on the economic dispatch and risk management of the large-scale plug-in electric vehicles (PHEVs)-penetrated power systems. *Renewable and Sustainable Energy Reviews* **16**(3) 1508–1515.

Plug-In Cars. 2014. Detailed list of electric cars and plug-in hybrids. Last accessed on May 20, 2014 at `http://www.plugincars.com/cars`.

Puterman, ML. 2005. *Markov decision processes: Discrete stochastic dynamic programming*. 1st ed. John Wiley & Sons.

Salas, DF, WB Powell. 2013. Benchmarking a scalable approximate dynamic programming algorithm for stochastic control of multidimensional energy storage problems. Tech. rep., Department of Operations Research and Financial Engineering, Princeton, NJ.

Scott, WR, WB Powell, S Moazehi. 2014. Least squares policy iteration with instrumental variables vs. direct policy search: Comparison against optimal benchmarks using energy storage. Tech. rep., Department of Operations Research and Financial Engineering, Princeton University.

Sioshansi, R, P Denholm. 2010. The value of plug-in hybrid electric vehicles as grid resources. *The Energy Journal* **31**(3) 1–23.

Sioshansi, R, SH Madaeni, P Denholm. 2014. A dynamic programming approach to estimate the capacity value of energy storage. *IEEE Transactions on Power Systems* **29**(1) 395–403.

Tang, X, N Liu, J Zhang, S Deng. 2012. Capacity optimization configuration of electric vehicle battery exchange stations containing photovoltaic power generation. *2012 7th International Power Electronics and Motion Control Conference*. Harbin, China, 2061–2065.

Tesla motors. 2014a. Battery swap. Last accessed on May 20, 2015 at `http://www.teslamotors.com/batteryswap`.

Tesla motors. 2014b. Battery swap pilot program. Last accessed on May 20, 2015 at `http://www.teslamotors.com/blog/battery-swap-pilot-program`.

Tesla motors. 2014c. Model X, Utility Meet Performance. Last accessed on May 20, 2015 at `http://www.teslamotors.com/modelx`.

Tesla motors. 2014d. Specs. Last accessed on May 20, 2015 at `http://www.teslamotors.com/models/specs`.

Tesla motors. 2015. Supercharger. Last accessed on May 20, 2015 at `http://www.teslamotors.com/supercharger`.

U.S. Department of Energy. 2014. Secretary moniz announces nearly $50 million to advance high-tech, fuel efficient american autos. Last accessed on May 20, 2014 at `http://energy.gov/articles/secretary-moniz-announces-nearly-50-million-advance-high-tech-fuel-efficient-american-autos`.

Wang, J, C Liu, D Ton, Y Zhou, J Kim, A Vyas. 2011. Impact of plug-in hybrid electric vehicles on power systems with demand response and wind power. *Energy Policy* **39**(7) 4016–4021.

Worley, O, D Klabjan. 2011. Optimization of battery charging and purchasing at electric vehicle battery swap stations. *2011 IEEE Vehicle Power and Propulsion Conference*. Chicago, IL, 1–4.

Yin, KK, H Liu, NE Johnson. 2002. Markovian inventory policy with application to the paper industry. *Computers & Chemical Engineering* **26**(10) 1399–1413.

Zhang, C-H, J-S Meng, Y-Z Cao, X Cao, Q Huang, Q-C Zhong. 2012. The adequacy model and analysis of swapping battery requirement for electric vehicles. *2012 IEEE Power and Energy Society General Meeting*. San Diego, CA, 1–5.

Zhang, D, WL Cooper. 2005. Revenue management for parallel flights with customer-choice behavior. *Operations Research* **53**(3) 415–431.

Appendix A: Proof of Theorem 1

Herein, we provide the full details of the proof for Theorem 1. First, we outline three lemmas that are utilized in the proof of Theorem 1.

LEMMA 1. *The function* $g_t(k|s_t,a_t) =$

$$\sum_{j=\max\{a_t^+ +1,k\}}^{s_t+a_t} p_{s_t+a_t-j} + \left[\sum_{i=s_t+a_t-j}^{\infty} p_i\right]_{\substack{j\geq k\\ j=a_t^+}}. \tag{20}$$

Proof.

$$g_t(k|s_t,a_t) = \sum_{j\in\{S|j\geq k\}} p_t(j|s_t,a_t) \tag{21}$$

$$= \sum_{\substack{j\geq k\\ a_t^+<j\leq s_t+a_t}} p_{s_t+a_t-j} + \left[q_{s_t+a_t-j}\right]_{\substack{j\geq k\\ j=a_t^+}} \tag{22}$$

$$= \sum_{j=\max\{a_t^+ +1,k\}}^{s_t+a_t} p_{s_t+a_t-j} + \left[\sum_{i=s_t+a_t-j}^{\infty} p_i\right]_{\substack{j\geq k\\ j=a_t^+}} \tag{23}$$

□

LEMMA 2. *The following two summations are equivalent.*

$$\sum_{j=k}^{s_t+a_t} p_{s_t+a_t-j} = \sum_{i=0}^{s_t+a_t-k} p_i \tag{24}$$

Proof.

$$\sum_{j=k}^{s_t+a_t} p_{s_t+a_t-j} = p_{s_t+a_t-k} + p_{s_t+a_t-(k+1)} + \ldots + p_{s_t+a_t-(s_t+a_t)} \tag{25}$$

$$= p_{s_t+a_t-k} + p_{s_t+a_t-(k+1)} + \ldots + p_0 = \sum_{i=0}^{s_t+a_t-k} p_i \tag{26}$$

□

LEMMA 3. *The following two summations are equivalent.*

$$\sum_{j=a_t^+ +1}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i = \sum_{i=0}^{\infty} p_i \tag{27}$$

Proof.

$$\sum_{j=a_t^+ +1}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i \tag{28}$$

$$= p_{s_t+a_t-(a_t^+ +1)} + p_{s_t+a_t-(a_t^+ +2)} + \ldots + p_{s_t+a_t-(s_t+a_t)} + \sum_{i=s_t-a_t^-}^{\infty} p_i \tag{29}$$

$$= p_{s_t-a_t^- -1} + p_{s_t-a_t^- -2} + \ldots + p_0 + \sum_{i=s_t-a_t^-}^{\infty} p_i \tag{30}$$

$$= \sum_{i=0}^{s_t-a_t^- -1} p_i + \sum_{i=s_t-a_t^-}^{\infty} p_i = \sum_{i=0}^{\infty} p_i \tag{31}$$

□

Utilizing these, we prove Theorem 1: *There exists optimal decision rules $d_t^* : s_t \to A_{s_t}$ for the EV-SSMP which are nonincreasing in s_t for $t = 1, \ldots, N-1$ when demand D_t is governed by a nonincreasing discrete distribution.*

Proof. The claim is shown by demonstrating that the EV-SSMP exhibits the following 5 conditions (Puterman 2005).

1. $r_t(s_t, a_t)$ is nondecreasing in s_t for all $a_t \in A'$.

   That $r_t(s_t, a_t)$ is nondecreasing in s_t for a fixed a_t means that for a fixed action (i.e., number of batteries charged or discharged), the expected immediate reward will be greater when the number of full batteries is greater. This coincides with intuition as more batteries can be swapped or discharged when there are more full batteries available thereby leading to more reward. Consider $s_t \geq \tilde{s}_t$, using $s_t + a_t - s_{t+1} = \min\{D_t, s_t - a_t^-\}$ for any value which D_t can assume, it suffices to show that

$$r_t(s_t, a_t) \geq r_t(\tilde{s}_t, a_t), \tag{32}$$

   using the expected immediate reward function

$$r_t(s_t, a_t) = \sum_{j=0}^{\infty} \Big[P(D_t = j)\big(\rho \min\{j, s_t - a_t^-\}\big) \Big] - K_t a_t^+ + J_t a_t^-. \tag{33}$$

   It suffices to show that

$$r_t(s_t, a_t) \geq r_t(\tilde{s}_t, a_t) \Leftrightarrow \tag{34}$$

$$\sum_{j=0}^{\infty} \Big[P(D_t = j)\big(\rho \min\{j, s_t - a_t^-\}\big) \Big] - K_t a_t^+ + J_t a_t^-$$

$$\geq \sum_{j=0}^{\infty} \Big[P(D_t = j)\big(\rho \min\{j, \tilde{s}_t - a_t^-\}\big) \Big] - K_t a_t^+ + J_t a_t^- \Leftrightarrow \tag{35}$$

$$\sum_{j=0}^{\infty} \Big[P(D_t = j)\big(\rho \min\{j, s_t - a_t^-\}\big) \Big] \geq \sum_{j=0}^{\infty} \Big[P(D_t = j)\big(\rho \min\{j, \tilde{s}_t - a_t^-\}\big) \Big]. \tag{36}$$

   Therefore, because $P(D_t = j)\rho$ is multiplied by both sides of the inequality in Equation (36) for all values of j, Equation (32) can be demonstrated by proving

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\}, \tag{37}$$

   for all possible values of j. Using a proof by cases, the three possible cases of demand $D_t = j$ with respect to $s_t - a_t^-$ and $\tilde{s}_t - a_t^-$ are considered: (a) $j \leq \tilde{s}_t - a_t^-$, $j \leq s_t - a_t^-$, (b) $j \geq \tilde{s}_t - a_t^-$, $j \leq s_t - a_t^-$, and (c) $j \geq \tilde{s}_t - a_t^-$, $j \geq s_t - a_t^-$. The case where j is greater than $s_t - a_t^-$ and less than $\tilde{s}_t - a_t^-$ does not need to be considered as it is not possible because $s_t \geq \tilde{s}_t$. In each case, Equation (37) is reduced to a valid statement.

   (a) $j \leq \tilde{s}_t - a_t^-$, $j \leq s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow j = j \tag{38}$$

(b) $j \geq \tilde{s}_t - a_t^-$, $j \leq s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow j \geq \tilde{s}_t - a_t^- \tag{39}$$

(c) $j \geq \tilde{s}_t - a_t^-$, $j \geq s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} \geq \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow s_t - a_t^- \geq \tilde{s}_t - a_t^- \Leftrightarrow s_t \geq \tilde{s}_t \tag{40}$$

2. $g_t(k|s_t, a_t)$ is nondecreasing in s_t for all $k \in S$ and $a_t \in A'$.

That $g_t(k|s_t, a_t)$ is nondecreasing in s_t for a fixed a_t and k means that the probability that the number of full batteries in the next state is greater than some threshold k is higher when the number of full batteries in the current state is greater. Consider $s_t \geq \tilde{s}_t$, it suffices to show that

$$g_t(k|s_t, a_t) \geq g_t(k|\tilde{s}_t, a_t) \Leftrightarrow \tag{41}$$

$$\sum_{j \in \{S|j \geq k\}} p_t(j|s_t, a_t) \geq \sum_{j \in \{S|j \geq k\}} p_t(j|\tilde{s}_t, a_t) \Leftrightarrow \tag{42}$$

$$\sum_{\substack{j \geq k \\ a_t^+ < j \leq s_t + a_t}} p_{s_t + a_t - j} + \left[q_{s_t + a_t - j} \right]_{\substack{j \geq k \\ j = a_t^+}} \geq$$

$$\sum_{\substack{j \geq k \\ a_t^+ < j \leq \tilde{s}_t + a_t}} p_{\tilde{s}_t + a_t - j} + \left[q_{\tilde{s}_t + a_t - j} \right]_{\substack{j \geq k \\ j = a_t^+}} \Leftrightarrow \tag{43}$$

$$\sum_{j = \max\{a_t^+ + 1, k\}}^{s_t + a_t} p_{s_t + a_t - j} + \left[\sum_{i = s_t + a_t - j}^{\infty} p_i \right]_{\substack{j \geq k \\ j = a_t^+}} \geq$$

$$\sum_{j = \max\{a_t^+ + 1, k\}}^{\tilde{s}_t + a_t} p_{\tilde{s}_t + a_t - j} + \left[\sum_{i = \tilde{s}_t + a_t - j}^{\infty} p_i \right]_{\substack{j \geq k \\ j = a_t^+}}. \tag{44}$$

Using a proof by cases, all cases of k with respect to a_t are considered. For each case, Equation (44) is reduced to a valid statement. Note that the second term of both the left hand side and right hand side of Equation (44) is only included when both $j \geq k$ and $j = a_t^+$, which represents when demand meets or exceeds supply. It is indicated in each case of the proof which of the terms are included in the summation based on the relationship between k and a_t.

(a) $a_t^+ \geq k \Rightarrow a_t^+ + 1 > k$

The second term of each summation appears as both $j \geq k$ and $j = a_t^+$ are satisfied. Using Lemma 3, Equation (45) is reduced to Equation (46).

$$\sum_{j = a_t^+ + 1}^{s_t + a_t} p_{s_t + a_t - j} + \sum_{i = s_t + a_t - a_t^+}^{\infty} p_i \geq \sum_{j = a_t^+ + 1}^{\tilde{s}_t + a_t} p_{\tilde{s}_t + a_t - j} + \sum_{i = \tilde{s}_t + a_t - a_t^+}^{\infty} p_i \Leftrightarrow \tag{45}$$

$$\sum_{i=0}^{\infty} p_i = \sum_{i=0}^{\infty} p_i \tag{46}$$

(b) $a_t^+ < k \Rightarrow a_t^+ + 1 \geq k$

The second term of each summation does not appear as $j = a^+$ will never be satisfied. Starting from Equation (44), Lemma 2 is utilized to arrive at a known valid statement.

$$\sum_{j=k}^{s_t + a_t} p_{s_t + a_t - j} \geq \sum_{j=k}^{\tilde{s}_t + a_t} p_{\tilde{s}_t + a_t - j} \Leftrightarrow \tag{47}$$

$$\sum_{i=0}^{s_t+a_t-k} p_i \geq \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \tag{48}$$

$$\sum_{i=0}^{\tilde{s}_t+a_t-k} p_i + \sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i \geq \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \tag{49}$$

$$\sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i \geq 0 \tag{50}$$

3. $r_t(s_t, a_t)$ is a subadditive function on $S \times A'$.

The subadditivity of $r_t(s_t, a_t)$ implies that the incremental effect on the expected total reward of charging less batteries (or discharging more batteries) is less when the number of full batteries is greater. Consider $a_t \geq \tilde{a}_t$ and $s_t \geq \tilde{s}_t$, using $s_t + a_t - s_{t+1} = \min\{D_t, s_t - a_t^-\}$ for any value which D_t can assume, it suffices to show that

$$r_t(s_t, a_t) + r_t(\tilde{s}_t, \tilde{a}_t) \leq r_t(s_t, \tilde{a}_t) + r_t(\tilde{s}_t, a_t) \Leftrightarrow \tag{51}$$

$$\begin{aligned}
&\sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, s_t - a_t^-\}\big)\Big] - K_t a_t^+ + J_t a_t^- \\
&+\sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, \tilde{s}_t - \tilde{a}_t^-\}\big)\Big] - K_t \tilde{a}_t^+ + J_t \tilde{a}_t^- \\
&\leq \sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, s_t - \tilde{a}_t^-\}\big)\Big] - K_t \tilde{a}_t^+ + J_t \tilde{a}_t^- \\
&+\sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, \tilde{s}_t - a_t^-\}\big)\Big] - K_t a_t^+ + J_t a_t^- \Leftrightarrow
\end{aligned} \tag{52}$$

$$\begin{aligned}
&\sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, s_t - a_t^-\}\big)\Big] + \sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, \tilde{s}_t - \tilde{a}_t^-\}\big)\Big] \\
&\leq \sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, s_t - \tilde{a}_t^-\}\big)\Big] + \sum_{j=0}^{\infty}\Big[P(D_t = j)\big(\rho \min\{j, \tilde{s}_t - a_t^-\}\big)\Big].
\end{aligned} \tag{53}$$

Therefore, because $P(D_t = j)\rho$ is multiplied by all terms in Equation (53), it suffices to show that

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\}, \tag{54}$$

for all values of j. Using a proof by cases, every relevant case of a_t and $\tilde{a}_t$, and each scenario for demand $D_t = j$ with respect to $s_t - a_t^-$, $\tilde{s}_t - \tilde{a}_t^-$, $s_t - \tilde{a}_t^-$, $\tilde{s}_t - a_t^-$ are considered. The case where $\tilde{a}_t \leq 0$ and $a_t \geq 0$ is excluded as this is not possible from the definition of subadditivity that $a_t \geq \tilde{a}_t$. For each case, Equation (54) is reduced to a valid statement.

(a) $\tilde{a}_t \geq 0,\ a_t \geq 0 \Rightarrow \tilde{a}_t^- = a_t^- = 0$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{55}$$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t\} = \min\{j, s_t\} + \min\{j, \tilde{s}_t\} \tag{56}$$

(b) $\tilde{a}_t \leq 0,\ a_t \geq 0 \Rightarrow \tilde{a}_t^- \geq 0,\ a_t^- = 0$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \tag{57}$$

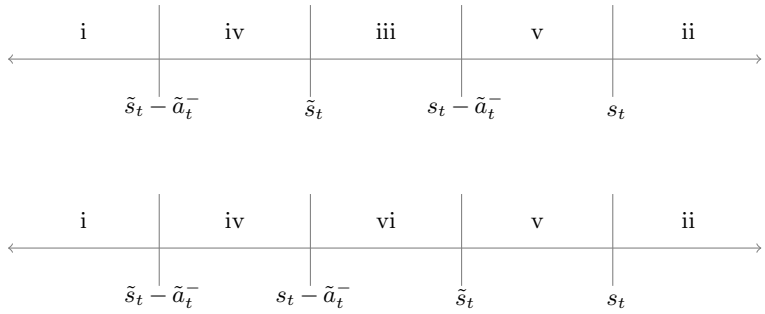


Figure 11 Scenarios of demand with respect to inventory for case (b).

Every possibility for demand j with respect to s_t, $\tilde{s}_t - \tilde{a}_t^-$, $s_t - \tilde{a}_t^-$, and $\tilde{s}_t$ is considered. Figure 11 is provided to aid the reader in visualizing the six possible scenarios. The ranges i-vi in the diagram correspond to the following scenarios i-vi.

i. $j \leq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq \tilde{s}_t,\ j \leq s_t - \tilde{a}_t^-,\ j \leq s_t$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \tag{58}$$

$$j + j \leq j + j \Leftrightarrow 2j = 2j \tag{59}$$

ii. $j \geq s_t \Rightarrow j \geq s_t - \tilde{a}_t^-,\ j \geq \tilde{s}_t,\ j \geq \tilde{s}_t - \tilde{a}_t^-$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \tag{60}$$

$$s_t + \tilde{s}_t - \tilde{a}_t^- = s_t - \tilde{a}_t^- + \tilde{s}_t \tag{61}$$

iii. $j \geq \tilde{s}_t,\ j \leq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-,\ j \leq s_t$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \tag{62}$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + \tilde{s}_t \Leftrightarrow \tilde{a}_t \geq 0 \tag{63}$$

iv. $j \leq \tilde{s}_t,\ j \leq s_t - \tilde{a}_t^-,\ j \geq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq s_t$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \tag{64}$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + j \Leftrightarrow \tilde{s}_t - \tilde{a}_t^- \leq j \tag{65}$$

v. $j \geq \tilde{s}_t,\ j \geq s_t - \tilde{a}_t^-,\ j \leq s_t \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \tag{66}$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + \tilde{s}_t \Leftrightarrow j \leq s_t \tag{67}$$

vi. $j \leq \tilde{s}_t,\ j \geq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-,\ j \leq s_t$

$$\min\{j, s_t\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\} \leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t\} \Leftrightarrow \tag{68}$$

$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + j \Leftrightarrow \tilde{s}_t \leq s_t \tag{69}$$

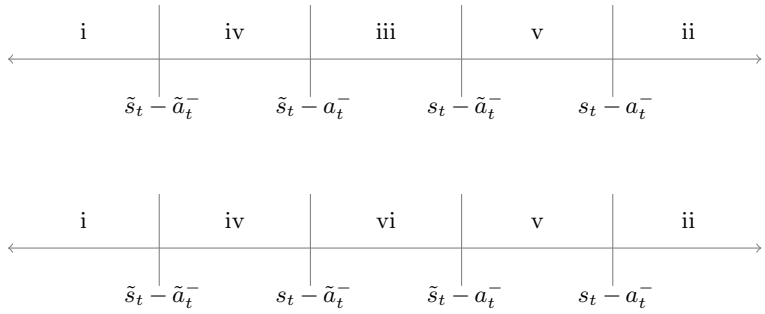


Figure 12 Scenarios of demand with respect to inventory for case (c).

(c) $\tilde{a}_t \leq 0,\ a_t \leq 0 \Rightarrow \tilde{a}_t^- \geq 0,\ a_t^- \geq 0,\ \tilde{a}_t^- \geq a_t^-$

Every possibility for demand j with respect to $s_t - a_t^-$, $\tilde{s}_t - \tilde{a}_t^-$, $s_t - \tilde{a}_t^-$, $\tilde{s}_t - a_t^-$ is considered. Figure 12 is provided to aid the reader in visualizing the six possible scenarios. The ranges i-vi in the diagram correspond to the following scenarios i-vi.

i. $j \leq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq \tilde{s}_t - a_t^-,\ j \leq s_t - \tilde{a}_t^-,\ j \leq s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$
$$\leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{70}$$
$$j + j \leq j + j \Leftrightarrow 2j = 2j \tag{71}$$

ii. $j \geq s_t - a_t^- \Rightarrow j \geq \tilde{s}_t - a_t^-,\ j \geq s_t - \tilde{a}_t^-,\ j \geq \tilde{s}_t - \tilde{a}_t^-$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$
$$\leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{72}$$
$$s_t - a_t^- + \tilde{s}_t - \tilde{a}_t^- = s_t - \tilde{a}_t^- + \tilde{s}_t - a_t^- \tag{73}$$

iii. $j \geq \tilde{s}_t - a_t^-,\ j \leq s_t - \tilde{a}_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-,\ j \leq s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$
$$\leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{74}$$
$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + \tilde{s}_t - a_t^- \Leftrightarrow \tilde{a}_t^- \geq a_t^- \tag{75}$$

iv. $j \leq \tilde{s}_t - a_t^-,\ j \leq s_t - \tilde{a}_t^-,\ j \geq \tilde{s}_t - \tilde{a}_t^- \Rightarrow j \leq s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$
$$\leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{76}$$
$$j + \tilde{s}_t - \tilde{a}_t^- \leq j + j \Leftrightarrow \tilde{s}_t - \tilde{a}_t^- \leq j \tag{77}$$

v. $j \geq \tilde{s}_t - a_t^-,\ j \geq s_t - \tilde{a}_t^-,\ j \leq s_t - a_t^- \Rightarrow j \geq \tilde{s}_t - \tilde{a}_t^-$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$
$$\leq \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{78}$$
$$j + \tilde{s}_t - \tilde{a}_t^- \leq s_t - \tilde{a}_t^- + \tilde{s}_t - a_t^- \Leftrightarrow j \leq s_t - a_t^- \tag{79}$$

vi. $j \le \tilde{s}_t - a_t^-$, $j \ge s_t - \tilde{a}_t^- \Rightarrow j \ge \tilde{s}_t - \tilde{a}_t^-$, $j \le s_t - a_t^-$

$$\min\{j, s_t - a_t^-\} + \min\{j, \tilde{s}_t - \tilde{a}_t^-\}$$
$$\le \min\{j, s_t - \tilde{a}_t^-\} + \min\{j, \tilde{s}_t - a_t^-\} \Leftrightarrow \tag{80}$$
$$j + \tilde{s}_t - \tilde{a}_t^- \le s_t - \tilde{a}_t^- + j \Leftrightarrow \tilde{s}_t \le s_t \tag{81}$$

4. $g_t(k|s_t, a_t)$ is a subadditive function on $S \times A'$ for all $k \in S$.

The subadditivity of $g_t(k|s_t, a_t)$ implies that the incremental effect of charging less batteries (or discharging more batteries) on the probability that the system moves to a state of full batteries above some threshold k is less when the number of full batteries is greater. Consider $a_t \ge \tilde{a}_t$ and $s_t \ge \tilde{s}_t$, it suffices to show that

$$g_t(k|s_t, a_t) + g_t(k|\tilde{s}_t, \tilde{a}_t) \le g_t(k|s_t, \tilde{a}_t) + g_t(k|\tilde{s}_t, a_t) \Leftrightarrow \tag{82}$$
$$\sum_{j=\max\{a_t^+ + 1, k\}}^{s_t + a_t} p_{s_t + a_t - j} + \left[\sum_{i=s_t + a_t - j}^{\infty} p_i\right]_{\substack{j \ge k \\ j = a_t^+}}$$
$$+ \sum_{j=\max\{\tilde{a}_t^+ + 1, k\}}^{\tilde{s}_t + \tilde{a}_t} p_{\tilde{s}_t + \tilde{a}_t - j} + \left[\sum_{i=\tilde{s}_t + \tilde{a}_t - j}^{\infty} p_i\right]_{\substack{j \ge k \\ j = \tilde{a}_t^+}}$$
$$\le \sum_{j=\max\{\tilde{a}_t^+ + 1, k\}}^{s_t + \tilde{a}_t} p_{s_t + \tilde{a}_t - j} + \left[\sum_{i=s_t + \tilde{a}_t - j}^{\infty} p_i\right]_{\substack{j \ge k \\ j = \tilde{a}_t^+}}$$
$$+ \sum_{j=\max\{a_t^+ + 1, k\}}^{\tilde{s}_t + a_t} p_{\tilde{s}_t + a_t - j} + \left[\sum_{i=\tilde{s}_t + a_t - j}^{\infty} p_i\right]_{\substack{j \ge k \\ j = a_t^+}}. \tag{83}$$

Using a proof by cases, every relevant case of k with respect to a_t and $\tilde{a}_t$ is considered. For each case, Equation (83) is reduced to a valid statement. The function $g_t(k|s_t, a_t)$ is comprised of two terms. The first term calculates the probability when demand never exceeds supply of batteries and the second calculates the probability that demand equals or exceeds supply. It is indicated in each case of the proof which of the terms are included in the summation based on the relationship between k, a_t, and $\tilde{a}_t$.

(a) $\tilde{a}_t^+ \ge k \Rightarrow a_t^+ \ge k$, $\tilde{a}_t^+ + 1 > k$, $a_t^+ + 1 > k$

For this case demand for battery swaps may exceed supply, therefore both terms of $g_t(k|s_t, a_t)$ appear.

$$\sum_{j=a_t^+ + 1}^{s_t + a_t} p_{s_t + a_t - j} + \sum_{i=s_t + a_t - a_t^+}^{\infty} p_i + \sum_{j=\tilde{a}_t^+ + 1}^{\tilde{s}_t + \tilde{a}_t} p_{\tilde{s}_t + \tilde{a}_t - j} + \sum_{i=\tilde{s}_t + \tilde{a}_t - \tilde{a}_t^+}^{\infty} p_i$$
$$\le \sum_{j=\tilde{a}_t^+ + 1}^{s_t + \tilde{a}_t} p_{s_t + \tilde{a}_t - j} + \sum_{i=s_t + \tilde{a}_t - \tilde{a}_t^+}^{\infty} p_i + \sum_{j=a_t^+ + 1}^{\tilde{s}_t + a_t} p_{\tilde{s}_t + a_t - j} + \sum_{i=\tilde{s}_t + a_t - a_t^+}^{\infty} p_i \Leftrightarrow \tag{84}$$
$$\sum_{i=0}^{\infty} p_i + \sum_{i=0}^{\infty} p_i \le \sum_{i=0}^{\infty} p_i + \sum_{i=0}^{\infty} p_i \Leftrightarrow \tag{85}$$
$$2\sum_{i=0}^{\infty} p_i = 2\sum_{i=0}^{\infty} p_i \tag{86}$$

(b) $\tilde{a}_t^+ < k$, $a_t^+ \geq k \Rightarrow \tilde{a}_t^+ + 1 \leq k$, $a_t^+ + 1 > k$

For this case, because $a_t^+ \geq k$, the second term of $g(k|s_t, a_t)$ does appear when action a_t is taken as demand can exceed supply. However, because $\tilde{a}_t^+ < k$, demand can never exceed supply when action $\tilde{a}_t$ is taken.

$$\sum_{j=a_t^+ +1}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{i=s_t+a_t-a_t^+}^{\infty} p_i + \sum_{j=k}^{\tilde{s}_t+\tilde{a}_t} p_{\tilde{s}_t+\tilde{a}_t-j}$$

$$\leq \sum_{j=k}^{s_t+\tilde{a}_t} p_{s_t+\tilde{a}_t-j} + \sum_{j=a_t^+ +1}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} + \sum_{i=\tilde{s}_t+a_t-a_t^+}^{\infty} p_i \Leftrightarrow \tag{87}$$

$$\sum_{i=0}^{\infty} p_i + \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{s_t+\tilde{a}_t-k} p_i + \sum_{i=0}^{\infty} p_i \Leftrightarrow \tag{88}$$

$$\sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{s_t+\tilde{a}_t-k} p_i \Leftrightarrow \tag{89}$$

$$\sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i + \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i \Leftrightarrow \tag{90}$$

$$0 \leq \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i \tag{91}$$

(c) $a_t^+ < k \Rightarrow \tilde{a}_t^+ < k$, $\tilde{a}_t^+ + 1 \leq k$, $a_t^+ + 1 \leq k$

For this case, demand for battery swaps never exceeds supply therefore, the second term of $g_t(k|s_t, a_t)$ does not appear when either action a_t or action $\tilde{a}_t$ are taken.

$$\sum_{j=k}^{s_t+a_t} p_{s_t+a_t-j} + \sum_{j=k}^{\tilde{s}_t+\tilde{a}_t} p_{\tilde{s}_t+\tilde{a}_t-j} \leq \sum_{j=k}^{s_t+\tilde{a}_t} p_{s_t+\tilde{a}_t-j} + \sum_{j=k}^{\tilde{s}_t+a_t} p_{\tilde{s}_t+a_t-j} \Leftrightarrow \tag{92}$$

$$\sum_{i=0}^{s_t+a_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i \leq \sum_{i=0}^{s_t+\tilde{a}_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \tag{93}$$

$$\sum_{i=0}^{\tilde{s}_t+a_t-k} p_i + \sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i$$

$$\leq \sum_{i=0}^{\tilde{s}_t+\tilde{a}_t-k} p_i + \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i + \sum_{i=0}^{\tilde{s}_t+a_t-k} p_i \Leftrightarrow \tag{94}$$

$$\sum_{i=\tilde{s}_t+a_t-k+1}^{s_t+a_t-k} p_i \leq \sum_{i=\tilde{s}_t+\tilde{a}_t-k+1}^{s_t+\tilde{a}_t-k} p_i. \tag{95}$$

In Equation (95) the number of terms on each side are exactly the same, however because $a_t \geq \tilde{a}_t$ the start of the summation is greater on the left hand side. Therefore, Equation (95) holds when $p_j = P(D_t = j)$ is governed by a nonincreasing discrete distribution.

5. $r_N(s_N)$ is nondecreasing in s_N.

Consider $s_N \geq \tilde{s}_N$, it suffices to show that $r_N(s_N) \geq r_N(\tilde{s}_N)$. This expression is reduced to a known valid statement.

$$r_N(s_N) \geq r_N(\tilde{s}_N) \Leftrightarrow \rho s_N \geq \rho \tilde{s}_N \Leftrightarrow s_N \geq \tilde{s}_N \tag{96}$$

□

Appendix B: Full Results from the Benchmark Policies

Table 3 Benchmark policy results for Winter.

					Backward Induction	Stationary Benchmark Policy			Dynamic Benchmark Policy		
Scenario	M	$\Phi(\%M)$	γ	ρ	Time (s)	Time (s)	Opt Gap	Demand Gap	Time (s)	Opt Gap	Demand Gap
1	104	0.42	5872	15.38	62.57	0.34	0.13	-0.01	0.34	0.13	-0.02
2	185	0.83	5359	21.54	77.02	14.11	0.11	0.00	13.95	0.14	0.00
3	88	0.48	1513	17.31	7.88	0.68	0.26	0.00	0.57	0.27	0.00
4	200	0.44	3821	12.69	51.27	11.75	0.48	-0.23	10.75	0.48	-0.23
5	158	0.58	1256	16.54	30.98	5.22	0.47	-0.01	5.67	0.50	-0.01
6	173	0.56	4077	24.62	46.9	7.95	0.11	0.00	7.55	0.13	0.00
7	123	0.62	1897	10.38	13.84	2.25	0.87	-0.42	2.67	0.88	-0.42
8	169	0.54	5615	15.00	44.41	6.97	0.19	-0.08	6.23	0.20	-0.08
9	73	0.79	1000	21.92	6.62	0.51	0.24	0.00	0.39	0.28	0.00
10	192	0.85	2667	23.85	84.32	16.31	0.22	0.00	14.79	0.26	0.00
11	162	0.27	2154	13.08	17.78	3.11	0.46	-0.13	2.32	0.46	-0.13
12	65	0.40	2795	22.69	5.17	0.21	0.07	0.03	0.2	0.06	0.01
13	119	0.50	4333	20.00	18.25	1.3	0.09	0.00	1.04	0.09	0.00
14	115	0.37	4590	10.00	9.46	0.97	1.02	-0.47	0.71	1.04	-0.49
15	96	0.87	1128	14.62	10.82	0.82	0.49	-0.14	0.73	0.53	-0.14
16	196	0.46	2410	20.77	52.97	12.07	0.26	0.00	10.35	0.28	0.00
17	146	0.38	6000	23.08	26.46	2.89	0.07	0.02	2.51	0.06	0.00
18	131	0.96	3051	13.46	26.49	3.67	0.41	-0.27	3.88	0.45	-0.27
19	177	0.94	4846	14.23	65.08	12.56	0.31	-0.18	12.56	0.34	-0.18
20	135	0.29	2538	22.31	14.73	1.64	0.13	0.00	1.36	0.14	0.00
21	54	0.73	4718	15.77	5.36	0.19	0.20	0.08	0.17	0.18	0.06
22	188	0.81	2026	12.31	61.4	14.79	0.67	-0.29	13.72	0.70	-0.29
23	154	0.90	1641	18.85	38.77	6.86	0.36	0.00	6.91	0.42	0.00
24	81	0.33	1769	10.77	4.36	0.32	0.74	-0.40	0.28	0.73	-0.40
25	138	0.98	3949	20.38	35.43	4.55	0.13	0.00	4.65	0.16	0.00
26	150	0.63	1385	23.46	29.37	4.78	0.29	0.00	4.31	0.33	0.00
27	181	0.35	4462	19.23	36.77	6.01	0.14	0.00	5.1	0.14	0.00
28	58	1.00	3308	16.92	6.1	0.26	0.17	0.09	0.24	0.14	0.02
29	142	0.67	4205	11.15	25.4	4.42	0.68	-0.39	3.86	0.69	-0.39
30	127	0.77	5487	17.69	29.55	3.11	0.12	0.03	2.85	0.11	0.00
31	100	0.75	2923	18.46	14.61	1.14	0.14	0.00	0.92	0.16	0.00
32	165	0.69	3436	18.08	62.34	7.68	0.22	0.00	7.47	0.24	0.00
33	92	0.88	5744	11.54	12.98	1.1	0.54	-0.24	1.06	0.55	-0.29
34	50	0.71	2282	11.92	3.85	0.16	0.46	-0.33	0.15	0.47	-0.37
35	108	0.65	3179	25.00	17.98	1.04	0.08	0.00	0.83	0.10	0.00
36	62	0.31	4974	19.62	5.03	0.2	0.05	0.02	0.18	0.16	0.07
37	77	0.60	5231	24.23	10.98	0.26	0.16	0.12	0.25	0.11	0.07
38	69	0.92	5103	21.15	10.52	0.23	0.17	0.12	0.22	0.14	0.09
39	85	0.52	3564	13.85	10.28	0.4	0.22	-0.15	0.24	0.21	-0.17
40	112	0.25	3692	16.15	11.95	0.38	0.11	0.00	0.35	0.13	0.02

Table 4 Benchmark policy results for Spring.

					Backward Induction	Stationary Benchmark Policy			Dynamic Benchmark Policy		
Scenario	M	$\Phi(\%M)$	γ	ρ	Time (s)	Time (s)	Opt Gap	Demand Gap	Time (s)	Opt Gap	Demand Gap
1	104	0.42	5872	5.51	58.32	0.49	0.11	0.05	0.83	0.07	0.04
2	185	0.83	5359	8.38	87.46	15.07	0.07	0.00	12.83	0.05	0.00
3	88	0.48	1513	6.41	9.76	0.61	0.16	0.00	0.61	0.12	0.00
4	200	0.44	3821	4.26	64.8	11.81	0.31	0.00	9.37	0.24	0.00
5	158	0.58	1256	6.05	39.51	5.49	0.32	0.00	4.49	0.27	0.00
6	173	0.56	4077	9.82	55.9	7.88	0.07	0.00	6.7	0.05	0.00
7	123	0.62	1897	3.18	18.69	2.21	0.76	-0.37	1.98	0.67	-0.37
8	169	0.54	5615	5.33	54.77	6.96	0.12	0.00	10.85	0.08	0.00
9	73	0.79	1000	8.56	8.46	0.55	0.14	0.00	0.83	0.11	0.00
10	192	0.85	2667	9.46	87.18	16.79	0.12	0.00	17.5	0.10	0.00
11	162	0.27	2154	4.44	23.66	3.29	0.33	0.00	3.71	0.28	0.00
12	65	0.40	2795	8.92	5.96	0.22	0.06	0.03	0.22	0.03	0.01
13	119	0.50	4333	7.67	21.06	0.93	0.06	0.00	1.29	0.04	0.00
14	115	0.37	4590	3.00	8.73	0.85	1.06	-0.60	0.68	0.92	-0.62
15	96	0.87	1128	5.15	14.11	0.84	0.31	0.00	0.83	0.26	0.00
16	196	0.46	2410	8.03	61.96	11.69	0.16	0.00	11.97	0.13	0.00
17	146	0.38	6000	9.10	31.09	2.96	0.06	0.02	3.05	0.03	0.01
18	131	0.96	3051	4.62	35.82	3.69	0.23	0.00	3.77	0.19	0.00
19	177	0.94	4846	4.97	81.21	12.25	0.17	0.00	14.82	0.14	0.00
20	135	0.29	2538	8.74	16.44	1.75	0.09	0.00	2.35	0.07	0.00
21	54	0.73	4718	5.69	5.99	0.18	0.18	0.11	0.26	0.12	0.08
22	188	0.81	2026	4.08	77.64	15.06	0.48	-0.01	16.03	0.41	-0.01
23	154	0.90	1641	7.13	47.9	6.78	0.22	0.00	7.08	0.18	0.00
24	81	0.33	1769	3.36	5.71	0.35	0.56	-0.21	0.34	0.45	-0.21
25	138	0.98	3949	7.85	42.05	4.57	0.07	0.00	4.34	0.06	0.00
26	150	0.63	1385	9.28	37.25	4.77	0.18	0.00	6.64	0.14	0.00
27	181	0.35	4462	7.31	43.57	6.02	0.09	0.00	6.27	0.06	0.00
28	58	1.00	3308	6.23	7.13	0.29	0.16	0.12	0.37	0.07	0.05
29	142	0.67	4205	3.54	34.85	4.27	0.41	-0.16	3.97	0.33	-0.16
30	127	0.77	5487	6.59	35.73	3.15	0.10	0.04	3.37	0.05	0.01
31	100	0.75	2923	6.95	17.49	1.03	0.08	0.00	1.14	0.06	0.00
32	165	0.69	3436	6.77	54.1	8.06	0.13	0.00	7.99	0.10	0.00
33	92	0.88	5744	3.72	16.95	1.07	0.26	0.09	1.19	0.18	0.02
34	50	0.71	2282	3.90	4.82	0.16	0.23	0.05	0.15	0.15	0.00
35	108	0.65	3179	10.00	18.95	0.96	0.05	0.00	1.15	0.03	0.00
36	62	0.31	4974	7.49	4.79	0.21	0.05	0.02	0.18	0.12	0.07
37	77	0.60	5231	9.64	10.85	0.25	0.15	0.12	0.44	0.08	0.06
38	69	0.92	5103	8.21	10.57	0.23	0.16	0.12	0.24	0.10	0.08
39	85	0.52	3564	4.79	11.48	0.28	0.14	0.03	0.27	0.08	0.00
40	112	0.25	3692	5.87	11.97	0.44	0.10	0.04	0.34	0.08	0.04

Table 5 Benchmark policy results for Summer.

					Backward Induction	Stationary Benchmark Policy			Dynamic Benchmark Policy		
Scenario	M	$\Phi(\%M)$	γ	ρ	Time (s)	Time (s)	Opt Gap	Demand Gap	Time (s)	Opt Gap	Demand Gap
1	104	0.42	5872	10.38	60.57	0.37	0.14	-0.03	0.35	0.11	-0.04
2	185	0.83	5359	16.54	79.45	15.1	0.08	0.00	13.54	0.07	0.00
3	88	0.48	1513	12.31	8.18	0.63	0.22	-0.03	0.52	0.18	-0.03
4	200	0.44	3821	7.69	54.13	12.15	0.49	-0.26	10.22	0.43	-0.26
5	158	0.58	1256	11.54	35.62	5.57	0.42	-0.04	5.07	0.36	-0.04
6	173	0.56	4077	19.62	50.69	7.98	0.09	0.00	7	0.07	0.00
7	123	0.62	1897	5.38	17.08	2.24	1.04	-0.33	2.17	0.95	-0.33
8	169	0.54	5615	10.00	45.74	7.48	0.19	-0.10	6.74	0.16	-0.10
9	73	0.79	1000	16.92	7.36	0.52	0.18	0.00	0.45	0.15	0.00
10	192	0.85	2667	18.85	82.94	17.26	0.16	0.00	15.01	0.13	0.00
11	162	0.27	2154	8.08	18.35	3.29	0.50	-0.15	2.34	0.44	-0.15
12	65	0.40	2795	17.69	5.02	0.21	0.07	0.02	0.2	0.03	0.01
13	119	0.50	4333	15.00	17.41	1.55	0.08	-0.02	0.93	0.05	-0.02
14	115	0.37	4590	5.00	9.31	0.77	1.44	-0.41	0.81	1.37	-0.43
15	96	0.87	1128	9.62	12.08	0.87	0.43	-0.10	0.71	0.38	-0.10
16	196	0.46	2410	15.77	57.96	11.79	0.21	0.00	9.72	0.18	0.00
17	146	0.38	6000	18.08	26.27	3.01	0.06	0.02	2.38	0.03	0.00
18	131	0.96	3051	8.46	28.91	3.73	0.39	-0.22	3.86	0.35	-0.22
19	177	0.94	4846	9.23	69.32	13.01	0.29	-0.16	12.43	0.25	-0.16
20	135	0.29	2538	17.31	13.73	1.81	0.12	0.00	1.27	0.09	0.00
21	54	0.73	4718	10.77	5.06	0.18	0.19	0.06	0.16	0.14	0.01
22	188	0.81	2026	7.31	68.47	15.07	0.66	-0.28	13.85	0.60	-0.28
23	154	0.90	1641	13.85	44.08	6.84	0.29	-0.02	7.29	0.25	-0.02
24	81	0.33	1769	5.77	4.25	0.33	0.93	-0.35	0.32	0.83	-0.35
25	138	0.98	3949	15.38	36.78	4.67	0.09	0.00	4.7	0.08	0.00
26	150	0.63	1385	18.46	32.77	5.01	0.23	0.00	4.58	0.19	0.00
27	181	0.35	4462	14.23	36.48	6.26	0.12	-0.02	5.06	0.09	-0.02
28	58	1.00	3308	11.92	5.96	0.24	0.16	0.06	0.2	0.09	-0.02
29	142	0.67	4205	6.15	27.89	4.21	0.80	-0.33	3.84	0.74	-0.33
30	127	0.77	5487	12.69	30.68	3.17	0.11	-0.02	2.85	0.07	-0.05
31	100	0.75	2923	13.46	14.82	1.02	0.11	-0.03	0.92	0.09	-0.03
32	165	0.69	3436	13.08	51.22	8.7	0.17	-0.03	7.62	0.14	-0.03
33	92	0.88	5744	6.54	11.96	1.07	0.57	-0.17	1.05	0.58	-0.24
34	50	0.71	2282	6.92	3.11	0.15	0.52	-0.25	0.14	0.49	-0.31
35	108	0.65	3179	20.00	16.1	0.91	0.06	0.00	0.74	0.05	0.00
36	62	0.31	4974	14.62	4.17	0.48	0.05	0.01	0.2	0.09	0.03
37	77	0.60	5231	19.23	9.28	0.25	0.15	0.12	0.36	0.08	0.06
38	69	0.92	5103	16.15	8.95	0.23	0.16	0.12	0.22	0.09	0.07
39	85	0.52	3564	8.85	8.52	0.28	0.24	-0.14	0.28	0.20	-0.17
40	112	0.25	3692	11.15	9.4	0.37	0.13	-0.03	0.34	0.10	-0.03

Table 6 Benchmark policy results for Fall.

					Backward Induction	Stationary Benchmark Policy			Dynamic Benchmark Policy		
Scenario	M	$\Phi(\%M)$	γ	ρ	Time (s)	Time (s)	Opt Gap	Demand Gap	Time (s)	Opt Gap	Demand Gap
1	104	0.42	5872	4.15	59.8	0.31	0.12	0.05	0.72	0.08	0.04
2	185	0.83	5359	6.62	78.79	13.8	0.08	0.00	13.81	0.06	0.00
3	88	0.48	1513	4.92	8.62	0.62	0.18	0.00	0.6	0.14	0.00
4	200	0.44	3821	3.08	60.82	11.03	0.38	0.00	10.65	0.30	0.00
5	158	0.58	1256	4.62	35.87	5.15	0.34	0.00	5.25	0.29	0.00
6	173	0.56	4077	7.85	50.78	8.5	0.08	0.00	6.94	0.05	0.00
7	123	0.62	1897	2.15	13.36	3.19	1.11	-0.71	2.02	0.98	-0.71
8	169	0.54	5615	4.00	48.23	6.64	0.14	0.00	6.28	0.11	0.00
9	73	0.79	1000	6.77	7.23	0.48	0.15	0.00	0.45	0.12	0.00
10	192	0.85	2667	7.54	80.56	16.02	0.14	0.00	15.71	0.11	0.00
11	162	0.27	2154	3.23	21.65	2.96	0.38	0.00	2.46	0.31	0.00
12	65	0.40	2795	7.08	5.15	0.21	0.07	0.03	0.2	0.04	0.01
13	119	0.50	4333	6.00	18.26	0.81	0.07	0.00	1.05	0.05	0.00
14	115	0.37	4590	2.00	5.12	0.87	2.23	-0.80	0.72	2.07	-0.81
15	96	0.87	1128	3.85	12.68	0.85	0.37	0.00	0.77	0.30	0.00
16	196	0.46	2410	6.31	58.1	12.52	0.17	0.00	10.29	0.13	0.00
17	146	0.38	6000	7.23	26.69	2.89	0.06	0.02	2.37	0.04	0.00
18	131	0.96	3051	3.38	31.55	3.66	0.29	0.00	3.71	0.24	0.00
19	177	0.94	4846	3.69	74.59	12.54	0.21	0.00	12.46	0.17	0.00
20	135	0.29	2538	6.92	14.81	1.73	0.09	0.00	1.23	0.07	0.00
21	54	0.73	4718	4.31	5.28	0.17	0.18	0.11	0.16	0.12	0.08
22	188	0.81	2026	2.92	71.83	15.52	0.59	-0.01	13.69	0.50	-0.01
23	154	0.90	1641	5.54	43.97	6.87	0.25	0.00	6.84	0.20	0.00
24	81	0.33	1769	2.31	3.6	0.32	0.94	-0.58	0.32	0.78	-0.58
25	138	0.98	3949	6.15	38.18	4.98	0.09	0.00	4.6	0.06	0.00
26	150	0.63	1385	7.38	33.18	4.91	0.19	0.00	4.24	0.15	0.00
27	181	0.35	4462	5.69	40.36	6.06	0.10	0.00	4.87	0.08	0.00
28	58	1.00	3308	4.77	6.25	0.28	0.16	0.12	0.28	0.08	0.05
29	142	0.67	4205	2.46	27.47	4.38	0.72	-0.45	3.91	0.60	-0.45
30	127	0.77	5487	5.08	32.26	3	0.10	0.04	2.86	0.05	0.01
31	100	0.75	2923	5.38	15.35	1.06	0.10	0.00	0.95	0.07	0.00
32	165	0.69	3436	5.23	48.88	7.97	0.15	0.00	7.31	0.12	0.00
33	92	0.88	5744	2.62	13.49	1.07	0.42	-0.08	1.02	0.34	-0.15
34	50	0.71	2282	2.77	3.9	0.15	0.34	-0.05	0.15	0.25	-0.10
35	108	0.65	3179	8.00	16.61	0.78	0.06	0.00	0.83	0.04	0.00
36	62	0.31	4974	5.85	4.29	4.89	0.05	0.02	0.32	0.12	0.06
37	77	0.60	5231	7.69	9.52	0.23	0.15	0.12	0.24	0.08	0.06
38	69	0.92	5103	6.46	9.37	0.21	0.17	0.12	0.24	0.09	0.07
39	85	0.52	3564	3.54	10.08	0.27	0.17	0.03	0.27	0.11	0.00
40	112	0.25	3692	4.46	10.7	0.34	0.11	0.04	0.35	0.09	0.04