

Practicality of Nested Risk Measures for Dynamic Electric Vehicle Charging

Daniel R. Jiang and Warren B. Powell

January 12, 2018

Abstract

We consider the sequential decision problem faced by the manager of an electric vehicle (EV) charging station, who aims to satisfy the charging demand of the customer while minimizing cost. Since the total time needed to charge the EV up to capacity is often *less* than the amount of time that the customer is away, there are opportunities to exploit electricity spot price variations within some reservation window. We formulate the problem as a finite horizon Markov decision process (MDP) and consider a risk-averse objective function by optimizing under a dynamic risk measure constructed using a convex combination of expected value and conditional value at risk (CVaR). It has been recognized that the objective function of a risk-averse MDP lacks a practical interpretation. Therefore, in both academic and industry practice, the dynamic risk measure objective is often not of primary interest; instead, the risk-averse MDP is used as a *computational tool* for solving problems with predefined “practical” risk and reward objectives (termed the *base model*). In this paper, we study the extent to which the two sides of this framework are compatible with each other for the EV setting — roughly speaking, does a “more risk-averse” MDP provide lower risk in the practical sense as well? In order to answer such a question, the effect of the degree of dynamic risk-aversion on the optimal MDP policy is analyzed. Based on these results, we also propose a principled approximation approach to finding an instance of the risk-averse MDP whose optimal policy behaves well under the practical objectives of the base model. Our numerical experiments suggest that EV charging stations can be operated at a significantly higher level of profitability if dynamic charging is adopted and a small amount of risk is tolerated.

1 Introduction

The recent popularity of electric vehicles (EVs) has spurred significant research attention in the area of EV storage management from a number of perspectives. One stream of literature, e.g., [Roe et al. \[2009\]](#), [Clement-Nyns et al. \[2010\]](#), [Sundstrom and Binding \[2010\]](#), [Sortomme et al. \[2011\]](#), [Rotering and Ilic \[2011\]](#), [Gan et al. \[2013\]](#), [Li et al. \[2014\]](#), [Qi et al. \[2014\]](#), takes an aggregate view of EVs and provides possible solutions to the negative impacts of an overloaded electrical infrastructure. Broadly speaking, the literature considers both coordinated and decentralized scheduling approaches to accomplish *peak-shifting*, i.e., flattening the load profile due to EV charging in order to reduce stress on the grid.

A complementary line of research deals with EV charging policies at the level of individual charging stations, focusing on the optimization of metrics such as charging cost and

quality of service (minimal blocking/waiting times in queue-based models). For example, Zhang et al. [2014] integrates a queueing approach with Markov decision process (MDP) theory to optimize mean waiting times under a cost constraint. In Bayram et al. [2011], Koutsopoulos et al. [2011], and Karbasioun et al. [2014], the authors propose optimal operating strategies for the setting where a charging station is paired with storage. The papers Bashash et al. [2010] and Bashash et al. [2011] use a detailed model of a lithium ion battery in order to optimize under the objective of cost of electricity and battery health degradation. The uncertainty in the use of EVs, i.e., the driving pattern of the customer, is taken into account in the charge management policy proposed by Iversen et al. [2014].

Motivation. The motivation behind the problem examined in this paper is new, but most closely related to the second stream of literature dealing with individual charging stations. Here, we consider the point of view of a firm that owns charging stations located in public areas, such as parking garages, gas stations, or hotels. The objective is to manage the charging process for a single charger and a single vehicle within the reservation window so that energy costs are low when subjected to stochastic spot prices. When the spot price is high, the manager has an incentive to delay charging until the price falls (at the risk of even higher prices in the future or the inability to fully charge the vehicle).

In order to reap the environmental and societal benefits that come with the widespread adoption of electric vehicles, a dense network of charging stations located nationwide is required, but such infrastructure is not possible without an economic incentive. Unfortunately, there is concern that such an incentive for owning and operating EV infrastructure may not exist in the current climate [Schroeder and Traber, 2012, Chang et al., 2012, Robinson et al., 2014]. Methods to make ownership of EV charging stations a more viable business model are thus a crucial area of study.

Reservation-based EV charging stations have recently grown in popularity as a way to increase efficiency and decrease range anxiety (i.e., the fear of being stranded away from a charging station with an empty battery). The implementation of reservation-based systems may have also stemmed from social and etiquette considerations: Caperello et al. [2013] suggest that the lack of well-understood behavioral guidelines (e.g., when is it acceptable to unplug another vehicle that is finished charging?) can inhibit the use of public charging stations due to the increased uncertainty. Through surveys, they find that many EV drivers prefer a reservation system so as to mitigate this uncertainty. Since 2011, reservation capabilities at public charging stations have been made available via smart-phone applications (see, e.g., ChargePoint, E-Charge HK in Hong Kong, and NextCharge). Moreover, the rise of the “sharing economy” has led to innovative peer-to-peer EV charging solutions such as Charge,¹ where users can make reservations for time on privately owned charging stations.

¹<https://www.chargie.net/>

Problem Overview. The typical scenario (see Figure 1) in our model includes the following sequence of events: (1) a customer reserves a charging window $[0, \tau]$ at some time $t < 0$, where τ can be longer than the time required for a full charge; (2) she drops off the EV at the charging station at $t = 0$; (3) the vehicle is (dynamically) charged during the interval $[0, \tau]$; (4) the customer returns at the end of the reservation window, with the expectation that the vehicle is adequately charged (if not, she given an *inconvenience compensation*). From the point of view of the overall model, the reservation length τ is random variable. Since the dynamic charging policy from $[0, \tau]$ is determined after τ is known, a family of optimal control problems parameterized by τ are eventually solved. The theoretical analysis of the control problem, however, is always for a given $\tau = T$.

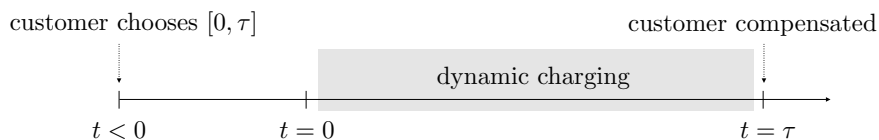


Figure 1: Sequence of Events

In a 2016 patent application by Google describing “mediator device for smart electric vehicle charging,” this exact situation is described as follows:

Modern EVSEs are equipped with technology that allows them to communicate with electric vehicles to control the charging rate. This capability allows the charging station to communicate the amount of charging capacity available to the EV... However, in many cases it is not necessary for the EV to charge the battery at the full rate. For example, EV drivers may drive their car[s] to work and leave [them] plugged in all day, even though charging completes in only a few hours. [Wytock et al., 2016]

We note that if the EV charging station is run on a first-come, first-served basis without reservations, our model could also be applicable if the smart-phone application asks the customer for an anticipated return time τ . The model is also flexible enough to exist within a larger strategy: for example, if the firm prefers to guarantee a full charge, then the interval $[0, \tau]$ could be split so that dynamic charging occurs first using our model, before switching to a static charging policy at the end.

The decision and risk analysis problem here is to balance the “risk” of the customer returning to an undercharged vehicle and the “reward” of completing a satisfactory charge at a reasonable cost. The extent to which this is an interesting problem depends largely on the power output of the charging station. For example, residential charging is typically quite slow and there is little opportunity to dynamically adjust the charging rate: Level 1 charging provides around 1.4 kW of output (a few miles per hour of charge) and Level 2 charging provides around 7.7 kW (or 20 miles per hour of charge).

However, technological improvements have resulted in the availability of high-powered chargers and it is now increasingly common for the customer to be away for a length of time that is *longer* than the amount of time needed to fully charge the EV. DC fast charging stations currently provide between 50 kW and 120 kW of power (see, e.g., Tesla’s Supercharger), providing 180 to 240 miles per hour of charge. As of 2017, a new tier of charging stations, billed as “ultra-fast” DC charging,²³ are slowly becoming available and can operate at a power rating of 350 to 400 kW. Given the high costs of the DC fast charging technology, estimated at \$65,000–\$70,000 compared to \$15,000–\$18,000 for a Level 2 charger [U.S. Department of Energy, 2012] (see also Bay Area Council [2011] and Schroeder and Traber [2012] for similar estimates), it is clear that optimal control policies can help ease the economics of such an investment. Indeed, our conversations with industry colleagues confirm that this problem is faced by owners of EV charging stations today.

The Importance of Risk. We suppose that the firm can directly interact with a spot energy market. The optimal policy depends on the spot prices of electricity P_t in the following way: if P_t is high, then we have a possible incentive to delay charging and buy later at a lower price. Of course, the trade-off here is that whenever charging is delayed, the risk of an inconvenienced customer (one who returns to a partially charged vehicle) increases. We assume in this paper that an inconvenienced customer is provided a financial compensation for her inconvenience when a vehicle is not charged adequately. The compensation includes a refund plus an additional spot price-dependent amount to cover the cost of purchasing the energy elsewhere. Therefore, the cost function used in our model is of the form

$$\text{total cost} = \text{cost of energy} + \text{inconvenience compensation},$$

which we emphasize is a *true financial cost* to the firm. Although a risk-neutral model (i.e., one that optimizes the expectation of the total cost) makes a trade-off between the “cost of energy” and the “inconvenience compensation” in order to minimize total cost, the following thought-experiment reveals that there are additional preferences that may not be captured under risk-neutrality: given two operating policies π_1 and π_2 that achieve the same expected total cost, where π_1 focuses on a low “inconvenience cost” at the expense of a high “cost of energy” and π_2 focuses on the opposite, which policy does one prefer? Since an inconvenienced customer may come with additional externalities (e.g., negative word of mouth, lost future sales) that are difficult to quantify, it is likely that π_1 is the preferred policy. This hints at the existence of *risk-aversion* toward events where the customer is inconvenienced.

Our proposed model uses the framework of risk-averse MDPs, introduced in Ruszczyński [2010], to capture the important issue of sequential risk in this problem. In this framework,

²<https://www.chargepoint.com/products/commercial/express-plus/>

³<https://www.engadget.com/2017/07/16/porsche-installs-first-super-fast-ev-chargers/>

the cost function remains unchanged: it continues to represent the *true financial cost* to the firm, as discussed above, but we replace the expected value operator with a *dynamic risk measure* [Frittelli and Gianin, 2004, Riedel, 2004, Pflug and Ruszczyński, 2005, Boda and Filar, 2006, Cheridito et al., 2006, Acciaio and Penner, 2011] in order to account for uncertain outcomes in a different way. When a *time-consistency* property holds for the dynamic risk measure (see Cheridito et al. [2006] and Ruszczyński [2010]), the risk-averse objective function for a finite horizon T becomes:

$$\min_{\pi \in \Pi} C_0^\pi + \rho_0(C_1^\pi + \rho_1(C_2^\pi + \cdots + \rho_{T-1}(C_T^\pi) \cdots)), \quad (1)$$

where Π is a set of policies, $\{C_t^\pi\}$ are costs incurred by following policy π , and $\{\rho_t\}$ are one-step coherent risk measures (i.e., components of the overall dynamic risk measure). A major computational advantage of this representation is that the familiar Bellman recursion applies for the value functions of the optimal risk-averse policy, allowing for established solution techniques to be adapted to the risk-averse setting. We provide a theoretical analysis of the *impact of risk-aversion* on the structure of the optimal policy for our EV charging problem. Related work to ours include risk-averse single and multi-stage newsvendor models [Bouakiz and Sobel, 1992, Eeckhoudt et al., 1995, Martínez-De-Albéniz and Simchi-Levi, 2006, Gotoh and Takano, 2007, Ahmed et al., 2007, Chen et al., 2009, Choi et al., 2011] and risk-averse joint-inventory and pricing models [Chen et al., 2007, Yang, 2017]. For single stage problems, the risk metrics used in these papers include conditional value at risk (CVaR) [Rockafellar and Uryasev, 2000], exponential utility, and mean-variance, while for multi-stage problems, they employ either additive utility functions or dynamic (coherent) risk measures. Choi et al. [2011] discuss the merits of *translation invariance* and *positive homogeneity*, two properties that our choice of methodology, dynamic coherent risk measures, possesses which the utility function framework does not.

Interpretability of the Risk-Averse Objective. In this paper, we address the issue of *interpretability* of (1) in a new way. It has been acknowledged that the main obstacle to the practicality of dynamic risk measures in MDPs is the lack of a clear economic interpretation [Rudloff et al., 2014, Iancu et al., 2015]. If $\rho_t = \mathbf{E}_t$ for all t , then (1) is equivalent to a quantity familiar to any practitioner, the expected total cost: $\min_{\pi} \mathbf{E}[\sum_t C_t^\pi]$. On the other hand, if ρ_t is chosen to be, say, CVaR, then (1) is evaluated with a nested version of CVaR [Cheridito and Stadje, 2009] and does not further simplify into a practical quantity. Nevertheless, the theoretical grounding and associated computational advantages of the framework [Ruszczyński, 2010, Philpott and de Matos, 2012, Philpott et al., 2013] makes it an attractive model for many: in fact, the Brazilian National System Operator (ONS) has *operationalized* a monthly hydrothermal planning strategy based on an MDP under nested mean-CVaR [Maceira et al., 2015]. To account for the interpretability issue, it is very common in both the literature and in practice to evaluate risk-averse policies generated by a risk-averse MDP along two dimensions, *practical risk* and *practical reward*, which may or

may not be directly related to the cost function of the MDP. See Figure 2 for an illustration of this approach. A few illustrative examples are listed below:

- [Philpott and de Matos \[2012\]](#) solve a hydro-thermal scheduling problem under a nested mean-CVaR risk measure. The practical reward is taken to be the (negative) “total operational cost” and the practical risk is the “number of violations where the national storage level is below a minimum threshold.” The authors explicitly acknowledge the discrepancy by stating that the dynamic risk measure objective “is arguably different from focusing on the distribution of the total annual cost, and controlling the extent of its upper tail.”
- [Shapiro et al. \[2013\]](#) optimize the Brazilian interconnected power system under a nested mean-CVaR risk measure, but effectiveness of the policy is gauged using the “expected cost” versus the “90%, 95%, and 99% quantiles of the cost.”
- [Çavus and Ruszczyński \[2014\]](#) compute risk-averse policies under a nested mean-semideviation risk measure for a credit limit decision problem and evaluates them by observing the profit distribution, where the implicit trade-off is between “expected profit” and “frequency of negative profits.”
- [Lin et al. \[2015\]](#) solve the [Almgren and Chriss \[2001\]](#) trade execution model under a general dynamic measure of risk. The numerical work evaluates the policies on the practical metrics of (negative) “mean cost” versus the “mean cost of the 30% worst samples.”
- [Maceira et al. \[2015\]](#) use a nested mean-CVaR for the hydro-planning MDP, but chooses the practical reward to be (negative) “expected value of operational cost” and practical risk to be “expected value of energy not supplied.”
- [Collado et al. \[2017\]](#) consider a stochastic path detection problem under nested mean-semideviation. The practical trade-off being considered is the “expected non-detection probability” versus the “variance of the non-detection probability.”

We emphasize that in each of these papers, the approach of Figure 2 is used. Although it is related, the dynamic risk measure objective used to find the policy does not match the practical objectives used for evaluation. Thus, we can view the risk-averse MDP as a *computational tool* for generating “risk-averse” policies that solve another model, which we term the *base model*. Based on trade-offs studied in the papers listed above, it is reasonable to specify a base model of the form

$$\text{maximize}_{\pi} [\text{Practical Reward}](\pi) \text{ subject to } [\text{Practical Risk}](\pi) \leq \varepsilon, \quad (2)$$

where π is an operating policy and ε is a parameter that controls the risk. Hence, the approach under discussion is when π is chosen to be the solution of a risk-averse MDP.

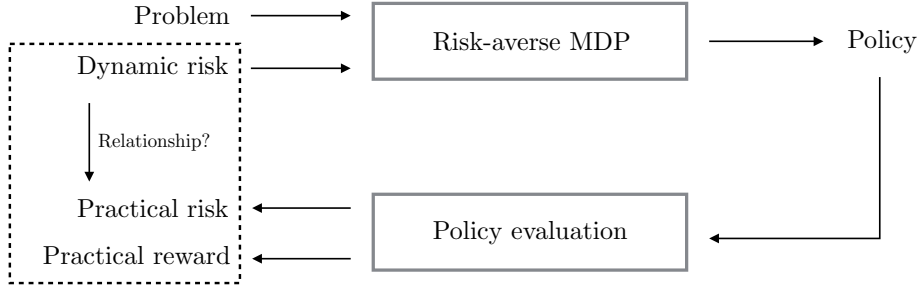


Figure 2: Evaluation using Practical Metrics of Risk and Reward

We note that this approach is typically not applied in a rigorous manner; the relationship between risk-aversion in the dynamic risk measure sense and the practical risk and reward is generally not well-understood. Our paper poses the question: does increased risk-aversion in the risk-averse MDP lead to lower practical risk in the base model as well? Such a property is desired as it indicates a *compatibility* between the risk-averse MDP and the practical objective. Without such a compatibility, the motivation for utilizing the risk-averse MDP model is quite weak. In this paper, we consider this question in the context of the dynamic EV charging problem along with several practical metrics that are of interest to industry. We show conditions under which compatibility holds for our model.

In current practice, the decision-maker is often limited by a computing budget and thus heuristically selects a few MDPs to solve. If one of the resulting policies generates an acceptable trade-off between practical risk and reward, then it is chosen for implementation; see, e.g., [Philpott and de Matos, 2012, Tables 6-7], [Shapiro et al., 2013, Figure 11], [Çavus and Ruszczyński, 2014, Table 8], and [Maceira et al., 2015, Figure 6] for examples of this general ad-hoc approach. Ideally, the decision-maker should investigate a larger set of risk-averse MDP policies. Along these lines, we offer a method of approximating the practical risk and reward for a spectrum of risk-averse policies using *polynomial optimization and sum of squares constraints* [Ahmadi and Majumdar, 2015]. The idea is that by solving a small number of risk-averse MDPs, our structural results allow us to reliably approximate the performance of an entire set of risk-averse optimal policies, providing practitioners an accessible means to explore a wider range of risk preferences. We then take advantage of the approximations to provide a systematic method for selecting a policy corresponding to a risk-averse MDP by analyzing the trade-offs of the practical metrics.

Our Results. The main contributions of this paper, some of which are of general interest beyond the EV application, are summarized below.

- We present a novel risk-averse, sequential model of the EV charging problem. The model is formulated using the framework of risk-averse MDPs with dynamic risk measures, constructed using one step risk measures that are a convex combination of expectation and conditional value at risk.

- We analyze the structural properties of the optimal policy. A main result here states that the charging thresholds of the optimal policy are nonincreasing in the current spot price (under a spot-price model incorporating seasonality, mean-reversion, and jumps) for any degree of dynamic risk-aversion.
- Related to the above results, we find conditions under which our risk-averse MDP is *compatible* (a notion that we will define) with a class of practical metrics of risk and reward when the approach of Figure 2 is adopted. This can be taken as justification that the use dynamic risk measures in risk-averse MDPs is an appropriate methodology in the case of the dynamic EV charging problem we present here.
- We propose an approximation strategy based on regression and polynomial optimization in order to guide the policy selection process, i.e., we address the question: how do we choose a risk-averse MDP such that the resulting optimal policy behaves well under the practical metrics?
- Finally, we demonstrate the proposed methods on a case study using spot price data from the California ISO (CAISO) and parking time distributions estimated from public garages in Santa Monica, California.

The paper is organized as follows. First, we present in Section 2 the mathematical model of the problem. In Section 3, we provide a structural analysis of the model. Next, in Section 4, we analyze the relationship of the policy to practical metrics of risk and reward and identify when our notion of compatibility holds. Lastly, the case study and the associated numerical work is presented in Section 6.

2 Mathematical Formulation

In this section, we first give a mathematical formulation of the dynamic EV charging problem. We then proceed to give a brief overview of dynamic risk measures in MDPs, along with the objective function used in the base model and objective used in the risk-averse MDP.

2.1 Dynamic EV Charging Model

Each realization T of the reservation length τ and each risk preference β defines a sequential decision problem of length $T + 1$, and in this paper, we are interested in the entire class of problems parameterized by T and β (hence, many relevant quantities will be indexed by both T and β). In this section and Section 3, we define and analyze the structure of the MDP under a fixed T and β . Next, in Section 4, we characterize the relationship between MDPs with varying β and also seek to understand the how the optimal policies behave

when evaluated using practical metrics of risk. The full distribution of τ is then taken into account in the numerical experiments of Section 6.

In our MDP, charging decisions are made at $t \in \{0, 1, \dots, T-1\}$, which are measured in units of 15 minutes. The time-based access fee to the charging station, paid by the customer per 15 minutes, is denoted c_f . The customer returns at time T and if applicable, receives an inconvenience compensation at time $T+1$ if an inadequate charge was provided (see Section 2.1.2). At every period t , our charging decision is a rate $x \in [0, x_{\max}]$ and using normalized units, we assume that x units of energy are obtained between t and $t+1$. Let $R_t \in [0, R_{\max}] =: \mathcal{R}$ be the amount of energy in storage at time t , where the *transition function* for R_t is given by $R_{t+1} = R_t + x_t$, provided that the charging decision x_t is chosen in the constraint set $\mathcal{X}(R_t) = \{x : 0 \leq x \leq \min\{R_{\max} - R_t, x_{\max}\}\}$.

Given our previous discussion on the wide range of charging station power ratings (1.4 kW to 400 kW), we note that it is increasingly likely that the maximum charge per period x_{\max} is larger than R_{\max} . This paper will explicitly consider two parameter regimes: (1) the *general charging regime* where x_{\max} is unrestricted and (2) the *fast charging regime* where x_{\max} is larger than R_{\max} . We note that x_{\max} depends on both the power output of the charging station and the length of a decision epoch; thus, the “fast charging regime” is a slight misnomer in that it arises when either the power output is high *or* the length of a decision epoch is large. The reason for the distinction between the two parameter regimes is that certain theorems (the results of Section 4) only hold in the fast charging case. Unless specifically noted, we consider the general charging regime, where no assumptions are made on x_{\max} .

2.1.1 Spot Price Dynamics

Now let us specify the dynamics of the spot price process $\{P_t\}$. Electricity prices exhibit several characteristics that should be captured: seasonality, mean-reversion, jumps, and potentially negative prices [Eydeland and Wolyniec, 2003, Kim and Powell, 2011, Zhou et al., 2015]. Because many conversations regarding dynamic EV charging technology are occurring in California, we obtained electricity price data from the California (CAISO) fifteen-minute market. These data, as shown in Figure 3, exhibits frequent negative prices but relatively light tails when compared to other markets (see, e.g., the five-minute NYISO real-time market).

Given these observations and the discrete-time decision problem, a reasonable choice to accommodate for the negative prices is to employ a discrete-time version of the jump-diffusion model of Cartea and Figueroa [2005] to directly fit the spot prices, rather than the log spot price. The continuous-time model is given by $P_t^c = g(t) + Y_t^c$. Here, $g(t)$ is a deterministic (sinusoidal) seasonality function and Y_t^c consists of a mean reverting component and a jump component; the dynamics are given by $dY_t^c = \kappa_Y (\mu_Y - Y_t^c) dt + \sigma_Y dW_t + J dN_t$, where $\kappa_Y > 0$ is the mean reversion parameter, θ_Y is the long term mean

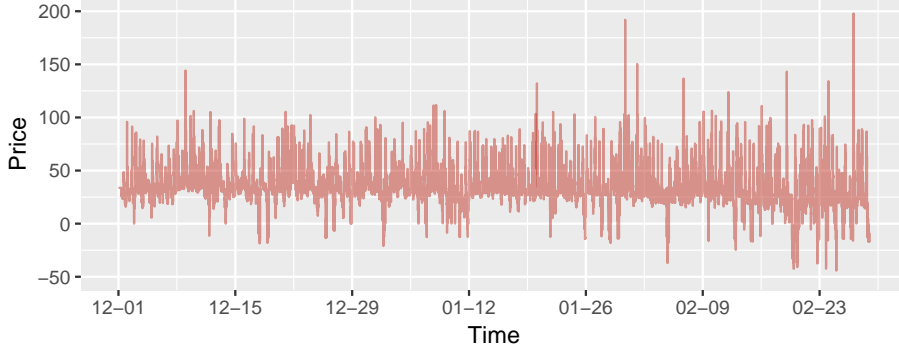


Figure 3: Historical Fifteen-Minute CAISO Prices, Winter 2016-2017

of the process, $\sigma_Y > 0$ is the volatility, W_t is a standard Brownian motion, J is a normally distributed jump size with mean μ_J and variance σ_J^2 , and N_t is a Poisson process with rate λ (all independent from one another).

The discrete-time approximation P_t of the continuous-time model at resolution Δt can be obtained by letting

$$P_t = g(t) + Y_t \quad \text{and} \quad Y_{t+\Delta t} = Y_t e^{-\kappa_Y \Delta t} + \mu_Y (1 - e^{-\kappa_Y \Delta t}) + \xi_{t+\Delta t} + X_{t+\Delta t} J_{t+\Delta t}, \quad (3)$$

where $\{\xi_t\}$ is a sequence of iid $\mathcal{N}(0, \sigma_Y^2 / (2\kappa_Y)(1 - e^{-2\kappa_Y \Delta t}))$ random variables, $\{X_t\}$ is a Bernoulli process with parameter $\lambda \Delta t$, and $\{J_t\}$ is a sequence of iid $\mathcal{N}(\mu_J, \sigma_J^2)$ jumps (a Bernoulli process approximation of the Poisson process, which assumes at most one jump per period). $\xi_{t+\Delta t}$, $X_{t+\Delta t}$, and $J_{t+\Delta t}$ are mutually independent. We utilize (3) with $\Delta t = 1$ for the remainder of this paper and assume that each time step Δt is 15 minutes. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the probability space on which all random variables are defined, where $\{\xi_t\}$, $\{X_t\}$, and $\{J_t\}$ are adapted to a filtration $\{\mathcal{F}_t\}_{t=0}^T$, with $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_T \subseteq \mathcal{F}$.

Note that by (3), we can define a random variable

$$\psi_{t+1} = g(t+1) - g(t) e^{-\kappa_Y} + \mu_Y (1 - e^{-\kappa_Y}) + \xi_{t+1} + X_{t+1} J_{t+1}, \quad (4)$$

which does not depend on the current spot price p , such that $p e^{-\kappa_Y} + \psi_{t+1}$ has the same distribution as $(P_{t+1} | P_t = p)$. Similarly, we define $\psi_{t+1,Y} = \psi_{t+1} - g(t+1)$, also independent of the current spot price p , such that $p e^{-\kappa_Y} + \psi_{t+1,Y}$ has the same distribution as $(Y_{t+1} | P_t = p)$. Both ψ_{t+1} and $\psi_{t+1,Y}$ are distributed as a mixture of normals. Throughout this paper, we use the notation $P_{t+1}(p)$ to represent a random variable equal in distribution $(P_{t+1} | P_t = p)$ and $Y_{t+1}(p)$ to represent $(Y_{t+1} | Y_t = p - g(t))$ in order to clarify certain expressions. Calibration details and the resulting model parameters are discussed in Section 6.

2.1.2 Customer Compensation

To properly assess the trade-offs in our model, we need to specify the compensation structure for a customer who returns to a vehicle with charge level R_T . Because our current optimal

control problem has only seen preliminary discussion [Wytock et al., 2016], there is no concrete, implemented example of how (and if) customers should be compensated in the case of a partial charge. Here, we present one reasonable choice that has arisen in discussion and remark that alternative structures satisfying similar convexity and monotonicity properties can be introduced into the model without significant difficulty.

Under a standard charging policy, the customer expects her vehicle to either be fully charged up to R_{\max} or has received a continuous charge for T periods, i.e., charged up to $R_0 + Tx_{\max}$. Therefore, the energy shortage is $h(R_T) = [\min(R_0 + Tx_{\max}, R_{\max}) - R_T]$. Let p_{ref} be a deterministic amount that represents the retail price per unit of energy, so that the customer is entitled to a refund of at least $h(R_T)p_{\text{ref}}$. The compensation also includes two additional payments, an *inconvenience amount* of γ_h per unit of shortage and a *market-dependent amount* of $\gamma_Y(P_{T+1} - g(T+1))$ per unit of shortage for some *positive, increasing* function γ_Y (see Assumption 1). The idea here is that if the shortage occurs in a time of high market prices (i.e., relative to the seasonality function), then the customer receives an additional payment to cover her costs at period $T+1$. The total inconvenience compensation at time $T+1$ is given by:

$$c_{T+1}(S_T, P_{T+1}) = [1 + \gamma_h h(R_T) + \gamma_Y(P_{T+1} - g(T+1))] h(R_T) p_{\text{ref}}.$$

The positivity of γ_Y is motivated by the notion that we should never reduce compensation payments to the customer, even in times of low or negative prices.

2.1.3 One-Period Cost Function

The *state variable* for this problem is $S_t = (R_t, P_t)$, which takes values in the state space $\mathcal{S} = \mathcal{R} \times \mathbb{R}$. We are now ready to define the cost function, which is given by $c_t(S_t, x_t) = (x_t P_t - c_f) \mathbf{1}_{\{t < T\}}$ for $t = 0, 1, \dots, T$ and the terminal cost is $c_{T+1}(S_T, P_{T+1})$, as defined above. We emphasize that both the cost of energy (for $t < T$) and the inconvenience compensation (for $t = T+1$) should be viewed as *actual financial costs* to the firm, and should not be interpreted to be artificial penalties intended to generate risk-averse policies. All risk preferences are added after the model of the problem is defined.

Let $\{x_0^\pi, x_1^\pi, \dots, x_{T-1}^\pi\}$ be a charging policy (indexed by π) where $x_t^\pi : \mathcal{S} \rightarrow [0, x_{\max}]$ is the decision function at time t . When following a policy π , we produce a sequence of states $S_0^\pi, S_1^\pi, \dots, S_T^\pi$ and obtain the costs $C_t^\pi = c_t(S_t^\pi, x_t^\pi(S_t^\pi))$, for $t = 0, 1, \dots, T$ and $C_{T+1}^\pi = c_{T+1}(S_T^\pi, P_{T+1})$ (note that $C_T^\pi = 0$). Let R_t^π be the level of charge at time t if a policy π is followed. Let $\mathbf{S}_T^\pi = (S_0^\pi, S_1^\pi, \dots, S_{T+1}^\pi)$ be the vector representing the sequence of states visited under policy π with $S_t^\pi = (R_t^\pi, P_t)$.

2.2 The Base Model & Practical Metrics

We now define a broad class of *practical reward* and *practical risk* metrics, along with a few motivating examples. Let f_T^+ be a function mapping a state vector for the problem of

horizon T into a “reward outcome,” so that the quantity $f_T^+(\mathbf{S}_T^\pi)$ is the random variable describing the practical reward of policy π . Similarly, define a function f_T^- so that $f_T^-(\mathbf{S}_T^\pi)$ is the practical risk of policy π . Finally, let r_{base} and ρ_{base} be monotone mappings from random variables into \mathbb{R} (e.g., \mathbf{E} or $\text{VaR}_{0.9}$) so that our metrics are given by

$$\text{Practical Reward} = r_{\text{base}}[f_\tau^+(\mathbf{S}_\tau^\pi)] \quad \text{and} \quad \text{Practical Risk} = \rho_{\text{base}}[f_\tau^-(\mathbf{S}_\tau^\pi)].$$

Note that we have replaced T with the random variable τ , in order to incorporate the uncertainty of the reservation horizon.

Although we have given a general definition above so that this framework can be applied in other application settings, in the EV charging problem, our primary goal is to use dynamic charging to reduce the cost of energy. Thus, the most sensible form of the practical reward is given by $r_{\text{base}} = \mathbf{E}$ and

$$f_\tau^+(\mathbf{S}_\tau^\pi) = c_f \tau - \sum_{t=1}^{\tau} (R_t^\pi - R_{t-1}^\pi) P_t - [1 + \gamma_h h(R_\tau^\pi) + \gamma_Y(Y_{\tau+1})] h(R_\tau^\pi) p_{\text{ref}}, \quad (5)$$

so that $r_{\text{base}}[f_\tau^+(\mathbf{S}_\tau^\pi)]$ represents the negative of the expected cost. On the other hand, any metric that we might reasonably want to minimize could be used for the practical risk. A few examples are given below.

- $\rho_{\text{base}} = \mathbf{E}$ and $f_\tau^-(\mathbf{S}_\tau^\pi) = [1 + \gamma_h h(R_\tau^\pi) + \gamma_Y(Y_{\tau+1})] h(R_\tau^\pi) p_{\text{ref}}$ gives the expected inconvenience compensation.
- Similarly, one could also consider letting $\rho_{\text{base}} = \text{VaR}_{0.9}$ and keeping f_τ^- as defined above so that the practical risk is the 90th quantile of the inconvenience compensation.
- $\rho_{\text{base}} = \mathbf{E}$ and $f_\tau^-(\mathbf{S}_\tau^\pi) = \mathbf{1}_{\{R_\tau/R_{\text{max}} \leq (1-\delta)\}}$ represents the probability that R_τ is less than some threshold of being full.
- $\rho_{\text{base}} = \text{CVaR}_{0.90}$ and $f_\tau^-(\mathbf{S}_\tau^\pi) = h(R_\tau^\pi)$ represents the average amount of shortage in the 10% worst cases.

Note that each of these risk metrics have a direct economic interpretation that can be useful in practice. We define the *base model* to be

$$\text{maximize}_\pi \quad r_{\text{base}}[f_\tau^+(\mathbf{S}_\tau^\pi)] \quad \text{subject to} \quad \rho_{\text{base}}[f_\tau^-(\mathbf{S}_\tau^\pi)] \leq \varepsilon, \quad (6)$$

for some “trade-off” parameter ε . Constrained MDPs, such as (6), are known to be very difficult to solve in general (see Altman [1999] for an overview), so many authors resort to using a risk-averse MDP to approximate the base model. To reiterate our previous discussions, we are interested in examining a notion of compatibility when policy evaluation is performed using (6) and policy optimization is done using a risk-averse MDP.

2.3 The Risk-Averse MDP Model

In this section, we discuss the risk-averse MDP model of [Ruszczynski \[2010\]](#), which serves as a surrogate to the computationally intractable base model given in (6). The eventual goal of this paper is to search over a parameterized class of MDPs in order to obtain an acceptable solution to (6).

The objective function of a risk-averse MDP is specified using a time-consistent dynamic risk measure $\rho_{0,T}$ of the *sequence* of downstream costs $C_0^\pi, C_1^\pi, \dots, C_{T+1}^\pi$ generated by a policy π . The objective is written using the notation $\rho_{0,T}(C_0^\pi, C_1^\pi, \dots, C_{T+1}^\pi)$; in the risk-neutral case, we would have $\rho_{0,T}(C_0^\pi, C_1^\pi, \dots, C_{T+1}^\pi) = \mathbf{E}\left[\sum_{t=0}^{T+1} C_t^\pi\right]$, but in general, we cannot assume such an additive form. The notion of time-consistency can be summarized as: if we prefer one stream of costs at some time in the future, and the costs from now until then are identical, then we must also prefer the same stream of costs *today*; see [Ruszczynski \[2010\]](#) for a precise definition. Let \mathcal{Z}_t be the space of \mathcal{F}_t measurable random variables. It can be shown, under some mild assumptions, that time-consistency results in the nested formulation

$$\rho_{0,T}(C_0^\pi, C_1^\pi, \dots, C_{T+1}^\pi) = C_0^\pi + \rho_0(C_1^\pi + \rho_1(C_2^\pi + \dots + \rho_T(C_{T+1}^\pi) \dots)), \quad (7)$$

for some *one-step conditional risk measures* $\rho_t : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t$.

The traditional risk-neutral optimization problem (see, e.g., [Puterman \[2014\]](#)) is

$$\min_{\pi \in \Pi} \mathbf{E} \left[\sum_{t=0}^{T+1} C_t^\pi \right] = \min_{\pi \in \Pi} C_0^\pi + \mathbf{E}_0(C_1^\pi + \dots + \mathbf{E}_T(C_{T+1}^\pi) \dots), \quad (8)$$

where \mathbf{E}_t is shorthand for $\mathbf{E}(\cdot | \mathcal{F}_t)$. We define the risk-averse optimization problem as $\min_{\pi \in \Pi} \rho_{0,T}(C_0^\pi, C_1^\pi, C_2^\pi, \dots, C_{T+1}^\pi)$, where the expected sum of costs is replaced with a dynamic risk measure $\rho_{0,T}$ acting on the sequence of costs. Applying the nested formulation of (7), we see that the risk-averse formulation is a clear analog of the risk-neutral case (8) with the conditional expectations replaced with one-step risk conditional risk measures:

$$\min_{\pi \in \Pi} C_0^\pi + \rho_0(C_1^\pi + \rho_1(C_2^\pi + \dots + \rho_{T-1}(C_T^\pi) \dots)). \quad (9)$$

In this paper, we follow several successful applications of risk-averse MDPs in the literature [[Philpott and de Matos, 2012](#), [Shapiro et al., 2013](#), [Maceira et al., 2015](#)] and take ρ_t to be the mean-CVaR risk measure. Every one-step conditional risk measure $\rho_t : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t$ has a “static” (as opposed to “dynamic”) counterpart that maps a random variable to \mathbb{R} , instead of another random variable. To avoid confusion, we skip directly to the static case.

Define the risk measures $\text{CVaR}_{\alpha_t} : \mathcal{Z} \rightarrow \mathbb{R}$ [[Rockafellar and Uryasev, 2000](#)] and $\rho_{\beta_t} : \mathcal{Z} \rightarrow \mathbb{R}$, where \mathcal{Z} is a space of random variables, to be

$$\begin{aligned} \text{CVaR}_{\alpha_t}(X) &= \inf_u \{u + (1 - \alpha_t)^{-1} \mathbf{E}[(X - u)^+]\}, \\ \rho_{\beta_t}(X) &= (1 - \lambda_t) \mathbf{E}(X) + \lambda_t \text{CVaR}_{\alpha_t}(X), \end{aligned} \quad (10)$$

where X is a random cost in \mathcal{Z} , $\lambda_t \in [0, 1]$ controls the emphasis on the tail risk, $\alpha_t \in (0, 1)$ controls the length of the tail, and $\beta_t = (\lambda_t, \alpha_t)$ is the *degree of risk-aversion* at time t . Note that ρ_{β_t} is a *coherent risk measure*, as axiomatized in Artzner et al. [1999], meaning that it satisfies the properties of convexity, monotonicity, translation invariance, and positive homogeneity. The *value at risk* for some risk level α_t of a random variable X is defined to be the α_t -quantile, or $\text{VaR}_{\alpha_t}(X) = \inf\{u : \mathbf{P}(X \leq u) > \alpha_t\}$.

Thus, we henceforth consider the optimization problem (9) with $\rho_t : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t$ taken to be the conditional counterpart to $\rho_{\beta_t} : \mathcal{Z} \rightarrow \mathbb{R}$; see Ruszczyński and Shapiro [2006], Ruszczyński [2010] for details and examples. Let $\beta = (\lambda_0, \alpha_0, \lambda_1, \alpha_1, \dots, \lambda_T, \alpha_T)$ control the *degree of dynamic risk-aversion* or *risk preference* of the MDP. The associated dynamic risk measure objective function is specified using β is written $\rho_{0,T}(C_0^\pi, C_1^\pi, \dots, C_{T+1}^\pi | \beta)$.

Let $\pi_T(\beta)$ be the (index of) an optimal policy for the decision problem of horizon $T + 1$ where ρ_t are specified using β . For convenience, we denote the decision function at time t for the optimal policy by $x_{t,T}(\cdot | \beta)$. The theorem below, proved in Ruszczyński [2010], serves as the foundation for risk-averse MDPs; it is an analog of the well-known risk-neutral Bellman recursion, with expectations replaced by one-step risk measures (in our case, they are taken to be ρ_{β_t}).

Theorem 1 (Risk-Averse Bellman Recursion, Ruszczyński [2010]).

Let the boundary condition be $V_{T,T}(s | \beta) = \rho_{\beta_T}[c_T(s, p e^{-\kappa_Y} + \psi_{T+1,Y})]$ for any state $s \in \mathcal{S}$. For each of the earlier time periods $t \in \{0, 1, \dots, T - 1\}$, let

$$V_{t,T}(s | \beta) = \min_{x \in \mathcal{X}(r)} c_t(s, x) + \rho_{\beta_t}[V_{t+1,T}(r + x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)].$$

The optimal policy to (9) is given by

$$x_{t,T}(s | \beta) \in \arg \min_{x \in \mathcal{X}(r)} c_t(s, x) + \rho_{\beta_t}[V_{t+1,T}(r + x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)],$$

i.e., it is greedy with respect to the optimal value function.

We now introduce some useful notation. First, it is often necessary for us to refer to a measure of risk on the future cost distribution; hence, we define:

$$\begin{aligned} v_{t,T}(r, p | \beta) &= \text{VaR}_{\alpha_t}[V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)], \\ c_{t,T}(r, p | \beta) &= \text{CVaR}_{\alpha_t}[V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)]. \end{aligned}$$

We will invoke the relationship

$$V_{t,T}(s | \beta) = \min_{x \in \mathcal{X}(r)} xp - c_f + \tilde{V}_{t,T}(r + x, p | \beta), \quad (11)$$

where $\tilde{V}_{t,T}(r + x, p | \beta) = \rho_{\beta_t}[V_{t+1,T}(r + x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)]$ is called the *post-decision value function*.

Before continuing, we point out that so far *two* optimization models, the base model (6) and the risk-averse MDP (9), have been defined. We re-emphasize that the relationship

between these two models is mostly explained in Figure 2, where (9) serves as a proxy for the more computationally challenging model (6). To be more precise, we are searching for a set of parameters β such that the solution to (9) is an acceptable solution to (6). This can be thought of as a *policy search* technique, wherein one searches over a subset of candidate policies; here, our candidate policies are solutions to different instances of the risk-averse MDP. As we reviewed in Section 1, this overall procedure is used by many authors (see, e.g., Philpott and de Matos [2012], Shapiro et al. [2013], Çavus and Ruszczyński [2014], Lin et al. [2015], Maceira et al. [2015], and Collado et al. [2017]). The reason that the risk-averse MDP (9) is often not taken to be the base model is due to its lack of practical interpretability. The exception is when $\rho_t = \mathbf{E}_t$, in which case (9) is a standard risk-neutral MDP whose objective is simply the expected cumulative cost. In the next section, we focus on analyzing (9) and in Section 4, we bring the models together and study a specific notion of compatibility between the two.

3 Structure of the Risk-Averse MDP

In this section, we state some structural properties of the risk-averse optimal policy and the associated optimal value function. All proofs that are not included in the body of the paper can be found in Appendix A.

Proposition 1 (Convexity). *Let $s = (r, p) \in \mathcal{S}$ be the state variable. The following statements hold for every t and β : (i) $x p + \tilde{V}_{t,T}(r+x, p | \beta)$ is convex in x on $\mathcal{X}(r)$, (ii) $\tilde{V}_{t,T}(s | \beta)$ is convex in the resource r , and (iii) $V_{t,T}(s | \beta)$ is convex in the resource r .*

The optimal value function also satisfies a monotonicity property, as shown in the following lemma.

Proposition 2 (Monotonicity). *For every t and resource level $r \in \mathcal{R} \setminus \{R_{\max}\}$, the optimal value function $V_{t,T}(r, p | \beta)$ is increasing in the spot price p .*

We now prove that the optimal policy is of the so called *basestock* or *order-up-to* type. In this result, the thresholds are given as a function of the risk parameters β and the spot price.

Theorem 2 (Optimal Policy). *For each time $t < T$ and spot price p , there exists a threshold resource level $r_{t,T}(p | \beta)$ such that the optimal policy takes the following form:*

$$x_{t,T}(r, p | \beta) = \min\{r_{t,T}(p | \beta) - r, x_{\max}\} \mathbf{1}_{\{r \leq r_{t,T}(p | \beta)\}},$$

where $r_{t,T}(p | \beta)$ is the smallest value of \tilde{r} that minimizes the quantity $\tilde{r} p + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$.

Sketch of Proof: The proof is standard and follows from convexity of the value function (i.e., Proposition 1) along with the fact that the risk measure is coherent (and thus, convex). \square

Our next goal is to analyze how p influences the threshold $r_{t,T}(p|\beta)$. We are able to show that if the compensation function γ_Y is “not too sensitive” to the spot price, then the charging thresholds are nonincreasing in the spot price (matching our intuition in that we should charge less when the prices are high). The sensitivity is formalized by the following Lipschitz condition.

Assumption 1. The market-dependent compensation function γ_Y is positive, increasing, and L_{γ_Y} -Lipschitz continuous where it holds that $L_{\gamma_Y} \leq p_{\text{ref}}^{-1} e^{2\kappa_Y}$.

Remark 1. If this assumption is not satisfied, then it is possible to observe strange (i.e., non-monotonic) behavior in the thresholds as a function of price. Note that the value of one unit of charge at the time of customer arrival is governed by γ_Y . Roughly speaking, the marginal cost we should consider is of the form “current price $-\gamma_Y$ (price deviation at $T+1$).” Assumption 1 limits the sensitivity of the marginal cost function contributed by γ_Y .

Before moving on to the theorem, we develop some notation and state a few technical lemmas.

Lemma 1 (Lipschitz in Spot Price). *If Assumption 1 holds, then for any t and r , the value function $V_{t,T}(r, \cdot | \beta)$ is Lipschitz continuous in the price dimension.*

Lemma 2 (Differentiability of Value at Risk). *For every resource state $r \in \mathcal{R} \setminus \{R_{\max}\}$, the value at risk function $v_{t,T}(r, p | \beta) = \text{VaR}_{\alpha_t}[V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)]$ is differentiable for almost every spot price p .*

Lemma 1 and Lemma 2 imply that both $V_{t,T}(r, \cdot | \beta)$ and $v_{t,T}(r, \cdot | \beta)$ are differentiable almost everywhere. It is reasonable for us to remove a null set of non-differentiability (call it \mathcal{V}) from the sample space under consideration for the remainder of the paper, i.e., those $\omega \in \Omega$ where $V_{t,T}(r, P_t(\omega) | \beta)$ or $v_{t,T}(r, \cdot | \beta)$ are not differentiable. Considering only $\Omega \setminus \mathcal{V}$ simplifies the discussion in the paper; it is reasonable to do so because $\mathbf{P}(P_t \in \mathcal{V}) = 0$ for any t (the price process P_t is a continuous random variable and exogenous to the decision process).

Lemma 3 (Differentiating under the Risk Measure). *Suppose the value function $V_{t+1,T}(\cdot | \beta)$ is differentiable at $r = r_0 \in \mathcal{R} \setminus \{R_{\max}\}$. Then, we may differentiate under the risk measure in the following sense:*

$$\begin{aligned} \partial_r \tilde{V}_{t,T}(r, p | \beta) |_{r=r_0} &= (1 - \lambda) \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)] |_{r=r_0} \\ &\quad + \lambda \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) | \psi_{t+1} \geq \text{VaR}_{\alpha_t}(\psi_{t+1})] |_{r=r_0}. \end{aligned}$$

With these lemmas in mind, we are able to show that the charging thresholds obey monotonicity in the spot price, as formalized in the theorem below. Similar theorems for the risk-neutral setting can be found in Secomandi [2010] and Chen et al. [2015]. Our analysis extends these results (although under a different MDP model) to the risk-averse setting by utilizing the preceding lemmas.

Theorem 3 (Nonincreasing Threshold in Spot Price). *Under Assumption 1, the threshold $r_{t,T}(p|\beta)$ is nonincreasing in the spot price p for any degree of dynamic risk-aversion β .*

Remark 2. Theorem 3 is proven here for a spot price model that contains mean-reversion, seasonality, and jumps, three properties that are most appropriate for electricity spot prices. However, we remark that the result can be adapted to other reasonable price models as well. The crucial property used in the proof of Theorem 3 is that $p - \mathbf{E}[P_{t+1} | P_t = p]$ is nondecreasing in p . For example, our framework can incorporate general models of the form $P_{t+1} = \zeta_{t+1}^1 P_t + \zeta_{t+1}^2$ for some random variables ζ_{t+1}^1 and ζ_{t+1}^2 , under the condition that $\mathbf{E}[\zeta_{t+1}^1] \leq 1$. Several popular choices, e.g., AR(1), martingale, and i.i.d. noise terms (along with any seasonality function), can be represented under this general model. Note that Assumption 1 is specialized for the model given in (3) and would need to be altered under a different spot price model.

4 Practical Risk Compatibility

The last section analyzes the risk-averse optimal policy under a *given* dynamic risk measure $\rho_{0,T}(\cdot|\beta)$. This is a standard approach in papers that consider risk-averse models; see, e.g., Lin et al. [2015], Devalkar et al. [2016], where the structure of the risk-averse optimal policy is explicitly characterized. However, in this section, we are interested in going one step further and commenting on the protocol of Figure 2 in the context of our EV charging problem. From this point of view, the risk-averse MDP model is simply a computational tool used to generate candidate policies that are then evaluated under a practical risk metric and a practical reward metric (the base model objectives). The motivation for doing so is that directly optimizing the practical risk and reward objectives can be computationally difficult, while the risk-averse MDP enjoys a dynamic programming decomposition.

4.1 Defining Risk Compatibility

When $\pi = \pi_T(\beta)$, we will use the convenient shorthand $\mathbf{S}_T^\beta = \mathbf{S}_T^\pi$ for the vector of visited states. We are interested in defining a notion of *compatibility* between the potentially divergent views of the problem. A new version of Figure 2 updated with mathematical notation is given in Figure 4. Before continuing, we need to first define the notion of “more risk-aversion.”

Definition 1 (Degree of Dynamic Risk-Aversion). Consider two dynamic risk measures $\rho_{0,T}$ and $\rho_{0,T}'$. If $\rho_{0,T}(\mathbf{C}) \leq \rho_{0,T}'(\mathbf{C})$ for any cost sequence \mathbf{C} , then $\rho_{0,T}'$ is *more risk-averse* than $\rho_{0,T}$.

To give a concrete example in the single period case, Belles-Sampera et al. [2014] points out that when CVaR is used instead of VaR in the financial industry (note that $\text{VaR}_\alpha(X) \leq$

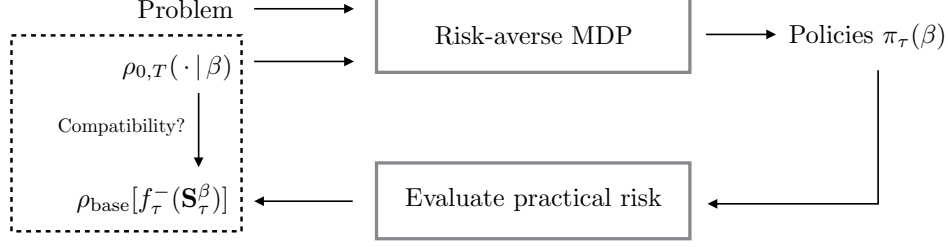


Figure 4: Evaluation using Practical Metrics of Risk and Reward

$\text{CVaR}_\alpha(X)$ for any cost X), then considerably higher capital reserves are required to offset the risk. It is thus reasonable to say that CVaR_α is more risk-averse than VaR_α .

Definition 2 (Practical Risk Compatibility). Let the dynamic risk measure be parameterized by β that varies within some set \mathcal{B} . Suppose that there is a partial order \leq on \mathcal{B} such that for $\beta \leq \beta'$, the dynamic risk measure $\rho_{0,T}(\cdot | \beta')$ is *more risk-averse* than $\rho_{0,T}(\cdot | \beta)$. We say that $\rho_{0,T}(\cdot | \beta)$ has the property of *practical compatibility on \mathcal{B} with respect to $f_T^-(\cdot)$* if the mapping $\beta \rightarrow f_T^-(\mathbf{S}_T^\beta)$ is nonincreasing for any T .

A consequence of this definition is that $\beta \rightarrow \rho_{\text{base}}[f_\tau^-(\mathbf{S}_\tau^\beta)]$ is nonincreasing. In other words, compatibility means that “more risk-aversion” in the dynamic risk measure sense should lead to lower risk in the practical sense as well. If, for example, more risk-aversion leads to higher practical risk, then we may conclude that the risk-averse MDP is incompatible with our practical objective and choose an alternative optimization procedure. The partial order for determining the degree of dynamic risk-aversion (in the sense of Definition 1) is generally easy to determine; for instance, $\text{CVaR}_\alpha(X)$ is interpreted as more risk-averse as α increases [Chen et al., 2009], and $(1 - \lambda) \mathbf{E}(X) + \lambda \text{CVaR}_\alpha(X)$ is more risk-averse as λ increases [Philpott and de Matos, 2012].

Note that the notion of compatibility depends on many factors: we must first define the dynamic risk measure $\rho_{0,T}(\cdot | \beta)$, the parameter set \mathcal{B} along with the partial order \leq , and the practical risk metric. Hence, for any given problem, we may observe compatibility in some cases and incompatibility in others (incompatibility refers to the situation when there are regions of \mathcal{B} when the mapping $\beta \rightarrow \rho_{\text{base}}[f_\tau^-(\mathbf{S}_\tau^\beta)]$ increases). An example can be found in Maceira et al. [2015], where it was observed that increasing risk-aversion in the dynamic risk measure unexpectedly *increased* the practical risk (energy not supplied):

It is possible to see that for some combinations of parameters $[\alpha \text{ and } \lambda]$ both operation cost and EENS [energy not supplied] increase simultaneously. Even though we expect higher operation costs when applying the CVaR mechanism, an increase in the EENS caused by preventive load curtailment is not desirable. Therefore, we call the region containing such combinations as “undesirable region.” On the other hand, there are combinations of parameters that provide

results located in a desirable area, i.e., cases where the operation cost increases and the EENS decreases.

Given that the system in Maceira et al. [2015] was operationalized in Brazil, the issue of compatibility is a critical one. Another example of incompatibility can be observed in Figure 11 of Shapiro et al. [2013]. A stylized illustration of these examples is given in Figure 5.

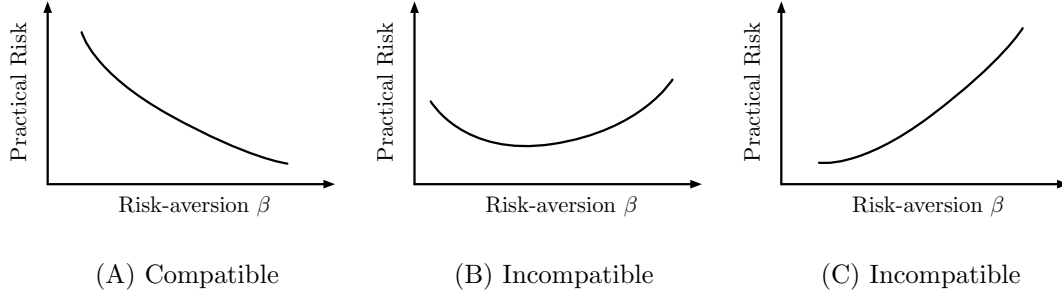


Figure 5: Illustration of Practical (In)compatibility

4.2 Compatibility Results

In the remainder of this section, we state and prove several risk compatibility properties under the condition that for any T , the function $f_T^-(\mathbf{S}_T^\pi)$ is componentwise nonincreasing in the resource states $\mathbf{R}_T^\pi = (R_0^\pi, R_1^\pi, \dots, R_T^\pi)$, a condition we refer to as *resource-monotonicity*. In other words, more energy in the vehicle corresponds to lower practical risk. Although we did not require resource-monotonicity when introducing practical risk metrics (in the interest of potential generalizations to other models), the inherent trade-offs associated with the “charging versus waiting” decision studied in the dynamic charging problem necessitate this property. We start with an important lemma.

Lemma 4 (Decreasing Marginal Value in Price). *Under Assumption 1, the pre-decision and post-decision marginal value of energy, $\partial_r V_{t,T}(r, p | \beta)$ and $\partial_r \tilde{V}_{t,T}(r, p | \beta)$, are nonincreasing in the spot price p for any fixed risk-aversion β .*

We remind the reader that in the proof of Theorem 3, we showed that $p + \partial_r \tilde{V}_{t,T}(r, p | \beta)$ is *nondecreasing* in p . The combined effect with the above lemma implies that $\partial_r \tilde{V}_{t,T}(r, p | \beta)$ cannot decrease too quickly, but is nonincreasing nonetheless. Leveraging this lemma, we are able to show our first compatibility result for the case when the dynamic risk measure is parameterized by $\beta_0 = (\lambda_0, \alpha_0)$. Recall that we described two parameter regimes, the general charging regime where x_{\max} is unrestricted and the fast charging regime where x_{\max} is larger than R_{\max} .

Theorem 4 (Compatibility in General Charging Regime). *Fix any reservation horizon T and the risk-aversion $\beta_{-0}^* = (\lambda_1^*, \alpha_1^*, \dots, \lambda_T^*, \alpha_T^*)$ for all periods except the first. Let*

$$\mathcal{B}(\beta_{-0}^*) = \{\beta = (\lambda_0, \alpha_0, \lambda_1^*, \alpha_1^*, \dots, \lambda_T^*, \alpha_T^*) : \lambda_0 \in [0, 1], \alpha_0 \in (0, 1)\}$$

be the set of risk-aversion parameters where (λ_t, α_t) matches $(\lambda_t^*, \alpha_t^*)$ of β_{-0}^* while the risk-aversion in the first period is a “free parameter.” Taking the partial order \leq to be the componentwise inequality on \mathbb{R}^2 gives compatibility on the set $\mathcal{B}(\beta_{-0}^*)$ between $\rho_{0,T}(\cdot | \beta)$ and any resource-monotone f_T^- .

We remark that the above compatibility result requires a very specific parameterization of the dynamic risk measure where only the initial risk-aversion parameters are free. This restrictive parameterization is not ideal in that it limits the decision-maker to a small set of risk-averse policies. A key takeaway from Theorem 4 and its narrow scope highlights the delicate nature of the practical risk compatibility question.

Interestingly, we now show in Theorem 5 that compatibility *can be shown with a fully general parameterization of the dynamic risk measure* if we focus only on the fast charging regime, where $R_{\max} \leq x_{\max}$. We include the proof of the theorem here in order to comment on the critical differences between the fast charging regime and the general regime described in Theorem 4 (see Section 4.3).

Theorem 5 (Compatibility in Fast Charging Regime). *Fix any reservation horizon T and consider the fully general risk-aversion parameter set*

$$\mathcal{B} = \{\beta : \lambda_t \in [0, 1], \alpha_t \in (0, 1), t = 0, 1, \dots, T\}$$

of risk-aversion parameters. Let the partial order \leq be the componentwise inequality on \mathbb{R}^T . Then, compatibility holds on the set \mathcal{B} between $\rho_{0,T}(\cdot | \beta)$ and any resource-monotone f_T^- .

Proof. Consider the risk-aversion parameters $\beta, \beta' \in \mathcal{B}$ where

$$\beta = (\lambda_0, \alpha_0, \dots, \lambda_T, \alpha_T) \leq (\lambda'_0, \alpha'_0, \dots, \lambda'_T, \alpha'_T) = \beta'.$$

Suppose that $r_{t,T}(p | \beta) \leq r_{t,T}(p | \beta')$ for any spot price p and time t . If this is true, then since the thresholds (i.e., charging “targets”) are higher for every t , it follows that $\mathbf{R}_T^\beta \leq \mathbf{R}_T^{\beta'}$. Therefore, threshold monotonicity $r_{t,T}(p | \beta) \leq r_{t,T}(p | \beta')$ is a sufficient condition for the compatibility criterion $f_T^-(\mathbf{S}_T^{\beta'}) \leq f_T^-(\mathbf{S}_T^\beta)$.

Our goal for the remainder of this proof is to show that the marginal value of energy *decreases* as risk-aversion increases: $\partial_r \tilde{V}_{t,T}(r, p | \beta) \geq \partial_r \tilde{V}_{t,T}(r, p | \beta')$. This would imply the desired condition $r_{t,T}(p | \beta) \leq r_{t,T}(p | \beta')$, from which the reasoning above can be applied to complete the argument.

Base Case. We proceed via backward induction, starting with the base case of $t = T - 1$. First, note that γ_Y is positive by Assumption 1 and ρ_{β_T} is nondecreasing in β_T . The derivative

$$\begin{aligned} \partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) \\ = -p_{\text{ref}} [1 + 2\gamma_h h(r) + \rho_{\beta_T} [\gamma_Y ((p e^{-\kappa_Y} + \psi_{t+1}) e^{-\kappa_Y} + \psi_{t+1,Y})]] \end{aligned}$$

is thus nonincreasing in β_T . Next, by Lemma 4 and the fact that $\text{VaR}_{\alpha_{T-1}}(\psi_T)$ is increasing in α_{T-1} , it follows that the mapping

$$\alpha_{T-1} \mapsto \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta) | \psi_T \geq \text{VaR}_{\alpha_{T-1}}(\psi_T)]$$

is nonincreasing. Hence, we have

$$\begin{aligned} \partial_r \tilde{V}_{T-1,T}(r, p | \beta) &= (1 - \lambda_{T-1}) \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta)] \\ &\quad + \lambda_{T-1} \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta) | \psi_T \geq \text{VaR}_{\alpha_{T-1}}(\psi_T)] \\ &\geq (1 - \lambda'_{T-1}) \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta')] \\ &\quad + \lambda'_{T-1} \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta') | \psi_T \geq \text{VaR}_{\alpha'_{T-1}}(\psi_T)] \\ &= \partial_r \tilde{V}_{T-1,T}(r, p | \beta'), \end{aligned}$$

completing the base case. The induction hypothesis is $\partial_r \tilde{V}_{t,T}(r, p | \beta) \geq \partial_r \tilde{V}_{t,T}(r, p | \beta')$ and we will verify that the same statement holds when $t - 1$ replaces t .

Inductive Step. Let $P = p e^{-\kappa_Y} + \psi_t$. We now consider the value function at time t evaluated at P (i.e., from the perspective of time $t - 1$):

$$V_{t,T}(r, p | \beta) = x_{t,T}(r, p | \beta) \cdot P - c_f + \tilde{V}_{t,T}(r + x_{t,T}(r, p | \beta), P | \beta), \quad (12)$$

and as a first step, we aim to show that $\partial_r V_{t,T}(r, p | \beta) \geq \partial_r V_{t,T}(r, P | \beta')$. Since we are in the fast charging regime, the optimal policy at time t is to either charge (all the way) up to the threshold $r_{t,T}(p | \beta)$ or do nothing if the charge level is already above $r_{t,T}(p | \beta)$. We define the price sets

$$\mathcal{P}_{\text{low}}(r, \beta) = \{P : r \leq r_{t,T}(p | \beta)\} \quad \text{and} \quad \mathcal{P}_{\text{high}}(r, \beta) = \{P : r_{t,T}(p | \beta) < r\}.$$

Due to $r_{t,T}(p | \beta)$ being nonincreasing in p , these price sets are disjoint and monotone in the sense that any element of $\mathcal{P}_{\text{low}}(r, \beta)$ is no greater than any element of $\mathcal{P}_{\text{high}}(r, \beta)$. Previously, in the proof of Theorem 3 (in Appendix A), the risk-aversion was fixed so the price sets did not need to explicitly depend on β ; however, now that we are studying the impact of risk-aversion, it is crucial to analyze this dependence. See Figure 6 for an illustration of the price sets: the two horizontal axes represent the price P and the vertical axis is the resource value r . The diagram shows that (by the induction hypothesis) the charging threshold $r_{t,T}(p | \beta)$ at t is nondecreasing as we move from β (the black curve) to β' (the red curve). This, in turn, alters the low and high price sets shown on the bottom horizontal axis for β (in black) and the top horizontal axis for β' (in red).

Doing casework on each of the spot price intervals, we can rewrite (12) for a fixed r as a piecewise function of P :

$$V_{t,T}(r, p | \beta) = \begin{cases} (r_{t,T}(p | \beta) - r) \cdot P - c_f + \tilde{V}_{t,T}(r_{t,T}(p | \beta), P | \beta) & \text{if } P \in \mathcal{P}_{\text{low}}(r, \beta), \\ -c_f + \tilde{V}_{t,T}(r, P | \beta) & \text{if } P \in \mathcal{P}_{\text{high}}(r, \beta). \end{cases}$$

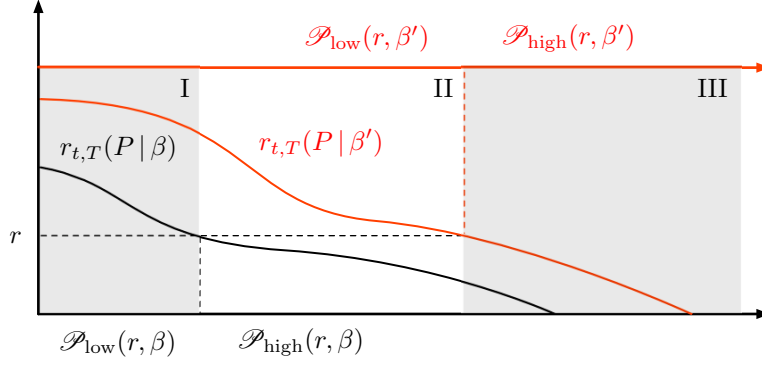


Figure 6: Illustration of the Price Sets for the Fast Charging Case

Differentiating in the resource r , we obtain

$$\partial_r V_{t,T}(r, p | \beta) = \begin{cases} -P & \text{if } P \in \mathcal{P}_{low}(r, \beta), \\ \partial_r \tilde{V}_{t,T}(r, p | \beta) & \text{if } P \in \mathcal{P}_{high}(r, \beta). \end{cases} \quad (13)$$

As shown in Figure 6, when comparing β and β' , there are three cases (I, II, and III) that we need to consider. In Case I, we have $P \in \mathcal{P}_{low}(r, \beta) \cap \mathcal{P}_{low}(r, \beta')$ and it is clear by (13) that $\partial_r V_{t,T}(r, p | \beta) = \partial_r V_{t,T}(r, P | \beta') = -P$. Case III is similarly straightforward because when $P \in \mathcal{P}_{high}(r, \beta) \cap \mathcal{P}_{high}(r, \beta')$, we know that by the induction hypothesis

$$\partial_r V_{t,T}(r, p | \beta) = \partial_r \tilde{V}_{t,T}(r, p | \beta) \geq \partial_r \tilde{V}_{t,T}(r, p | \beta') = \partial_r V_{t,T}(r, P | \beta').$$

Now let us move to Case II, when $P \in \mathcal{P}_{high}(r, \beta) \cap \mathcal{P}_{low}(r, \beta')$. Observe the following important property whenever $P \in \mathcal{P}_{high}(r, \beta) = \{P : r_{t,T}(p | \beta) < r\}$. As characterized in Theorem 2, the threshold $\tilde{r} = r_{t,T}(p | \beta)$ minimizes the quantity $\tilde{r}P + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$. By convexity of $\tilde{r}P + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$ in \tilde{r} (Proposition 1), it follows that the derivative $P + \partial_r \tilde{V}_{t,T}(\tilde{r}, p | \beta)$ evaluated at some $\tilde{r} > r_{t,T}(p | \beta)$ must be nonnegative. Hence, we have for Case II:

$$\partial_r V_{t,T}(r, p | \beta) = \partial_r \tilde{V}_{t,T}(r, p | \beta) \geq -P = \partial_r V_{t,T}(r, P | \beta')$$

which completes the argument that $\partial_r V_{t,T}(r, p | \beta) \geq \partial_r V_{t,T}(r, P | \beta')$ for any P . Because r was arbitrarily fixed, the inequality holds for any r as well. To finish the inductive step, i.e. $\partial_r \tilde{V}_{T-1,T}(r, p | \beta) \geq \partial_r \tilde{V}_{T-1,T}(r, p | \beta')$, the proof of the base case can be essentially repeated. \square

4.3 Incompatibility

What goes wrong in the slow charging case if we wish to employ the fully general \mathcal{B} from Theorem 5? As with logic introduced in the proof of Theorem 3 in Appendix A (see Figure 15), extending the analysis of Theorem 5 to the slow charging case also requires the

definition of *three* price sets: $\mathcal{P}_{\text{low}}(r, \beta) = \{P : r + x_{\max} < r_{t,T}(p | \beta)\}$, $\mathcal{P}_{\text{med}}(r, \beta) = \{P : r \leq r_{t,T}(p | \beta) \leq r + x_{\max}\}$, and $\mathcal{P}_{\text{high}}(r, \beta) = \{P : r_{t,T}(p | \beta) < r\}$.

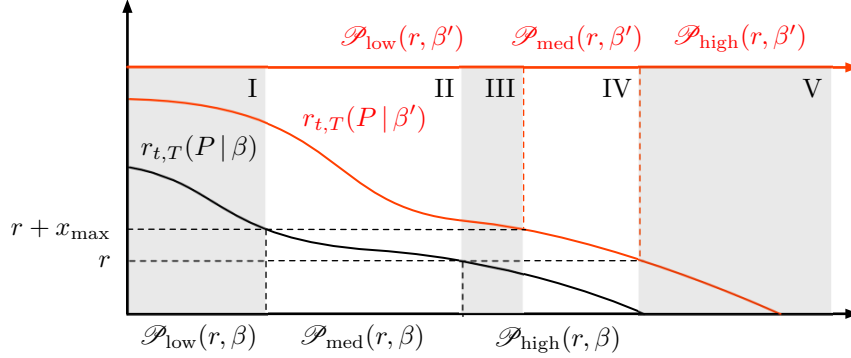


Figure 7: Illustration of the Price Sets for the Slow Charging Case

In Figure 7, we illustrate a situation where $\mathcal{P}_{\text{high}}(r, \beta) \cap \mathcal{P}_{\text{low}}(r, \beta') \neq \emptyset$; this region is denoted “III.” The two quantities that we need to compare in this region to show $\partial_r V_{t,T}(r, p | \beta) \geq \partial_r V_{t,T}(r, P | \beta')$ are $\partial_r \tilde{V}_{t,T}(r, p | \beta)$ and $\partial_r \tilde{V}_{t,T}(r + x_{\max}, P | \beta')$ — unfortunately, it is *not necessarily true* that the former is greater than the latter. Note, however, that Case I, II, IV, and V have the correct inequality. One can confirm via numerical experiments that incompatibility over \mathcal{B} in the slow charging case indeed exists.

5 Selecting the Degree of Dynamic Risk-Aversion

So far, we have studied the dynamic EV charging problem along two dimensions: (1) we provide an in-depth characterization on the structure of the optimal policy for a *fixed* degree of risk-aversion β in Section 3, and (2) we consider *varying* β and analyze the issue of compatibility between optimal policies of risk-averse MDPs (specified under a dynamic risk measure built by nesting mean-CVaR) and resource-monotone practical risk metrics in Section 4. Now that the approach shown in Figures 2 and 4 has been justified (to an extent) by the theorems of Section 4, one question still remains: *how do we select β so that the optimal risk-averse MDP policy performs “well” under the base model?*

As we previously discussed, the current practice is to heuristically select a large batch of potential β ’s, solve a large number of MDPs, simulate a large number of resulting policies, and then choose one that provides an acceptable trade-off (see, e.g., Maceira et al. [2015]). Unfortunately, this heuristic tuning method requires significant computational resources, negating the original motivation to solve practical reward versus risk problem via the simple and computationally effective dynamic programming approach (refer to Figure 2). In addition to CPU time, another weakness of the heuristic tuning approach is that the final choice of β must be from the pre-selected batch.

We now propose a simple approximation procedure that mitigates these issues. It utilizes structural information from our compatibility results so that β can be selected in a more principled manner. Suppose that we are given practical metrics $r_{\text{base}}[f_{\tau}^{+}(\mathbf{S}_{\tau}^{\pi})]$ and $\rho_{\text{base}}[f_{\tau}^{-}(\mathbf{S}_{\tau}^{\pi})]$, along with a set $\bar{\mathcal{B}} \subseteq \mathcal{B}$ and \leq that are compatible with f_{τ}^{-} for all T . Given a “trade-off parameter” ε , our goal is to search for some $\beta \in \bar{\mathcal{B}}$ such that the practical risk $\rho_{\text{base}}[f_{\tau}^{-}(\mathbf{S}_{\tau}^{\beta})] \leq \varepsilon$ while $r_{\text{base}}[f_{\tau}^{+}(\mathbf{S}_{\tau}^{\pi})]$ is maximized; see the formulation given in (6).

Step 1 (Solve MDPs). Sample a small set of risk-aversion parameters $\{\beta^1, \beta^2, \dots, \beta^n\}$ and solve the associated MDPs for each T . For every i , use simulation to obtain estimates $y^i \approx r_{\text{base}}[f_{\tau}^{+}(\mathbf{S}_{\tau}^{\beta^i})]$ and $z^i \approx \rho_{\text{base}}[f_{\tau}^{-}(\mathbf{S}_{\tau}^{\beta^i})]$.

Step 2 (Structured Regression). Utilize the “datasets” of practical reward/risk versus β

$$\{(\beta^1, y^1), (\beta^2, y^2), \dots, (\beta^n, y^n)\} \quad \text{and} \quad \{(\beta^1, z^1), (\beta^2, z^2), \dots, (\beta^n, z^n)\}$$

to construct regression models $\hat{r}_{\text{base}}(\beta)$ and $\hat{\rho}_{\text{base}}(\beta)$ that respectively approximate the practical reward and risk incurred by the MDP policies specified by β . Leverage the structural compatibility result to complement the regression procedure (see below).

Step 3 (Policy Selection). Use the models $\hat{r}_{\text{base}}(\beta)$ and $\hat{\rho}_{\text{base}}(\beta)$ to find a parameter $\hat{\beta}$ such that the practical reward $\hat{r}_{\text{base}}(\hat{\beta})$ is maximized while the practical risk is constrained: $\hat{\rho}_{\text{base}}(\hat{\beta}) \leq \varepsilon$. Solve the MDPs associated with $\hat{\beta}$ and implement the policy.

The main idea of Step 2 is to take advantage of Theorems 4 and 5 and employ a form of structured regression to ensure that $\hat{\rho}_{\text{base}}(\beta)$ is nonincreasing in β . Many monotone regression techniques are specified only for the one dimensional case [Mukerjee, 1988, Dette et al., 2006], but following the general idea given in Ahmadi and Majumdar [2014] for convex regression, if we restrict $\hat{\rho}_{\text{base}}$ to take the form of a polynomial, then the monotonicity constraints can be approximated with *sum of squares* (SOS) constraints. The resulting problem, for an l_1 error function, can be solved efficiently using semi-definite programming [Parrilo, 2003]. For example, suppose we are using the set $\bar{\mathcal{B}} = \{\beta : (\lambda, \alpha) = \beta_0 = \beta_1 = \dots = \beta_T\}$, the set of risk parameters that are time-independent. This set can be parameterized by two scalars, λ and α . Letting $\beta^i = (\lambda^i, \alpha^i)$, the structured regression problem is:

$$\begin{aligned} & \underset{\text{poly } \hat{\rho}_{\text{base}}}{\text{minimize}} && \sum_{i=1}^n |\hat{r}_{\text{base}}(\lambda^i, \alpha^i) - z^i| \\ & \text{subject to} && -\partial_{\lambda} \hat{r}_{\text{base}}(\lambda, \alpha), -\partial_{\alpha} \hat{r}_{\text{base}}(\lambda, \alpha) \text{ are SOS.} \end{aligned}$$

We illustrate this procedure using numerical experiments in Section 6.

6 Numerical Results and Case Study

In this section, we investigate our results on a realistic instance of the problem in the presence of a DC fast charging station (e.g., the ones described in Section 1 by Tesla and

ChargePoint). Specifically, we empirically verify the structural results of Theorem 2 and Theorem 3, in addition to the compatibility result (for the fast charging case) of Theorem 5. We then show the effectiveness of the risk-averse MDP selection procedure described in Section 5. The stochastic models of the price process $\{P_t\}$ and the reservation length τ are informed by market data from CAISO and parking garage inventory data from the city of Santa Monica, California.

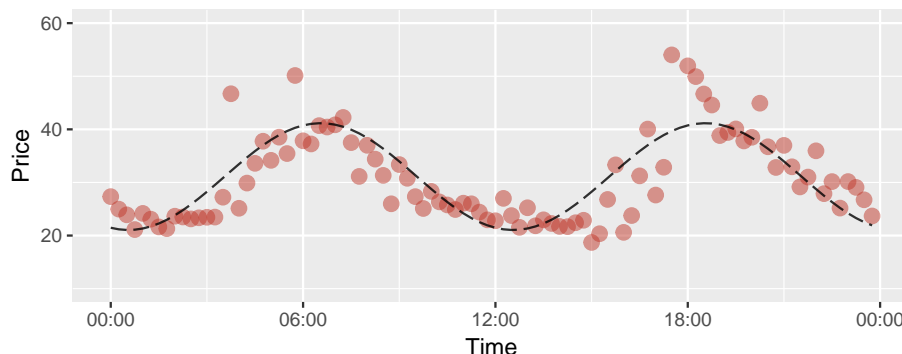


Figure 8: Seasonality Function $g(t)$ and Average (Across Time) Prices

6.1 Stochastic Model & Problem Parameters

The spot market in our model is based on a dataset from the CAISO fifteen minute market (the BAYSHOR2_1_N001 node) during the winter period December 1, 2016 to February 28, 2017. Since our problem is at the time scale of a few hours at most, we are primarily interested in understanding the *intra-day* seasonality. To accomplish this, for each 15 minute period during the day, we compute the average price across all weekdays in the dataset and then fit a sinusoidal function (with a period of one day) to the resulting time series; see Figure 8. The estimated seasonality function is given by

$$g(t) = 13.586 \sin(2\pi t/48) - 0.7597 \cos(2\pi t/48) + 34.1362,$$

where t measured in units of 15 minutes.

After removing the seasonality, we used maximum likelihood estimation to find the parameters of the Y_t stochastic process, resulting in $\mu_Y = -0.492$, $\kappa_Y = 0.341$, $\sigma_Y = 5.350$, $\mu_J = -0.484$, $\sigma_J = 40.602$, and $\lambda = 0.131$. Figure 9 provides a comparison between actual and fitted prices that includes both $g(t)$ and Y_t .

Although we do not have access to EV charging reservation data, we examined a large dataset⁴ of parking space availability in public parking garages of Santa Monica, California (updated one every five minutes). An analysis of *the changes in parking garage availability over time* across 14 parking garages and one year of five minute data from May 2016 to April 2017 (nearly 1.5 million records) provided us a estimated distribution for τ . Observations

⁴Santa Monica Open Data, <https://data.smgov.net/Transportation/Parking-Lot-Counts/ng8m-khuz>

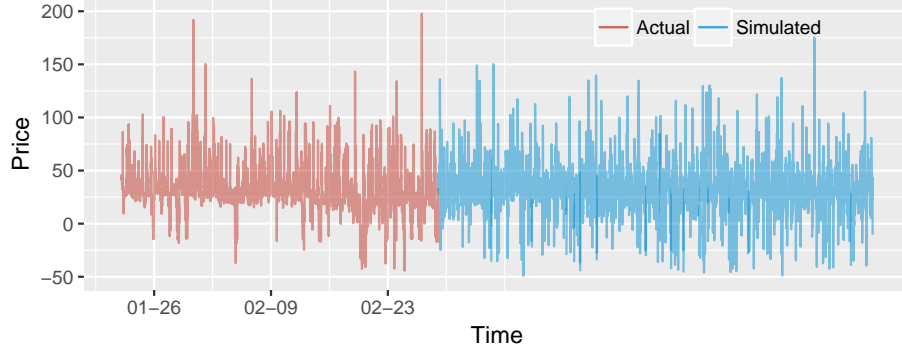


Figure 9: Actual and Simulated Spot Prices

of “parking time” were computed by first identifying times of peak utilization (i.e., the local maximums) and then analyzing the rate at which the garage empties. For example, if the garage empties quickly approximately one hour after the peak, then this suggests that many vehicles are parked for a length of one hour. The process yielded over 700,000 “parking observations” which were then used to estimate a distribution; see the second panel of Figure 10. The first panel of Figure 10 shows a one month sample of the raw data from the parking utilization dataset. It is interesting to note that a significant portion of parking garage patrons park for around 1.25 hours.

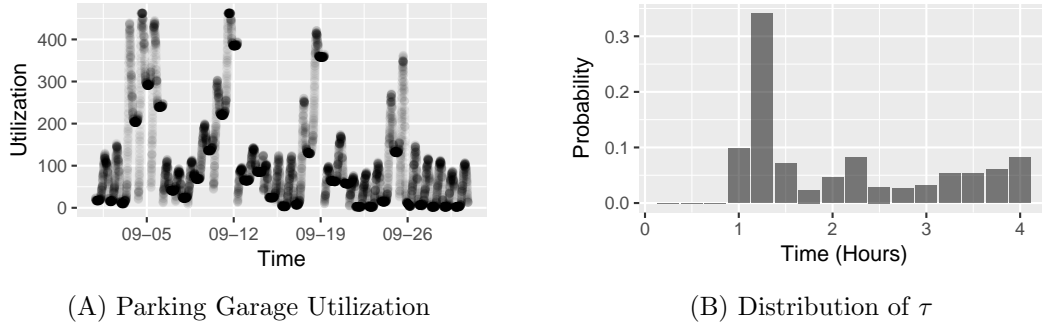


Figure 10: Raw Parking Garage Data & Estimated Distribution of Reservation Length τ

The initial price is chosen to be $P_0 = \$35/\text{MWh}$ (we follow the convention of wholesale electricity prices being written in units of MWh^{-1} even though our setting is on the order of kWh) and the battery is assumed to be initially empty: $R_0 = 0$. The capacity of the battery is $R_{\max} = 60 \text{ kWh}$, in line with the entry version of the Tesla Model S from 2012 to 2015. This is a reasonable choice for a medium sized battery, as there are both EVs with smaller capacities (e.g., Nissan Leaf or BMW i3) and EVs with larger capacities (e.g., newer and higher end Tesla models). We set the charging fee to be $c_f = \$2.00/\text{hour}$. The parameters related to the inconvenience compensation are taken to be $p_{\text{ref}} = \$0.05/\text{kWh}$, $\gamma_h = 0.01$, and $\gamma_Y(y) = \log(1 + \exp(y))$ (a smooth version of $\max(0, y)$), where y is in units of $\$/\text{kWh}$.

6.2 Exact MDP Results

In order to properly solve the risk-averse MDPs, discretization is necessary. We use the method of Roy [2003] to discretize ψ_{t+1} to only contain integer outcomes and we ignore the outcomes that occur with probability less than $1.5 \cdot 10^{-3}$. Such a discretization gives us 260 distinct spot prices allows us to utilize the linear programming method for computing the CVaR, as pointed out in Rockafellar and Uryasev [2000]. The resource state is also discretized so that R_t takes values in $\{0, 1, \dots, R_{\max}\}$. Each MDP contains approximately 16,000 states per time period, leading to 270,000 total states when $\tau = 16$.

We consider the risk-aversion set $\bar{\mathcal{B}} = \{\beta : (\lambda, \alpha) = \beta_0 = \beta_1 = \dots = \beta_T\}$. To adequately explore $\bar{\mathcal{B}}$, we solve a total of 4,953 MDPs for parameter values of $\lambda \in \{0, 0.05, \dots, 1\}$, $\alpha \in \{0.05, 0.1, \dots, 0.95\}$, and $T \in \{4, 5, \dots, 16\}$. Computing the solution to this set of MDPs requires a massive amount of computation and we do it only to illustrate the structural properties of the problem. This step is not necessary in practice, as we will see in Section 6.3 that the approximation approach described in Section 5 is quite effective.

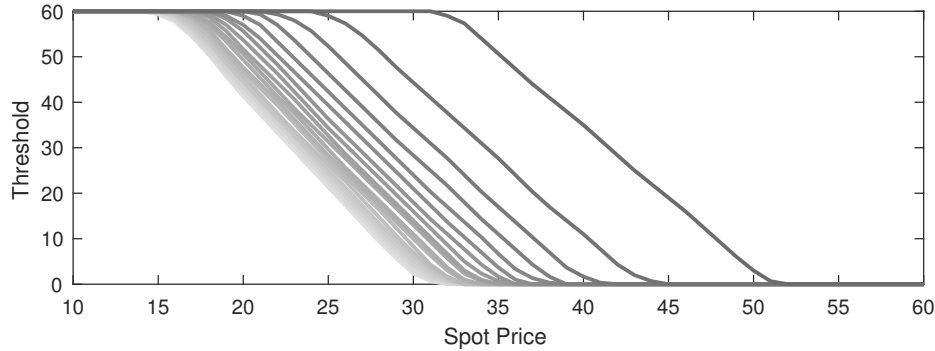


Figure 11: Thresholds for $\lambda = 0.5$, $T = 16$

We report our results using Figures 11 and 12. In Figure 11, we plot the thresholds as a function of price for a fixed $\lambda = 0.5$. Each curve represents a different value of α , which range from 0.05 to 0.95 (darker means higher α value). This serves to experimentally verify the statement of Theorem 3, that the thresholds are nonincreasing in p . Moreover, it verifies that the thresholds are nondecreasing in α , as shown in the proof of Theorem 5. We notice that at the for a given degree of risk-aversion, a \$30 change in spot price can alter the threshold by $R_{\max} = 60$.

Figure 12 fixes the spot price p and displays the variation of the threshold as a function of β (again, the thresholds are nondecreasing by Theorem 5) for the MDP associated with $T = 16$. We see that the degree of risk-aversion (λ, α) can dramatically change the policy: in the risk-neutral case for $p = 25$, the threshold $r_{t,T}(p | \beta) \approx 20$, while in the most risk-averse case, $r_{t,T}(p | \beta) \approx 60 = R_{\max}$ (meaning we should charge as much as possible). This observation emphasizes the importance of selecting an appropriate value for β , which we discuss in the next section.

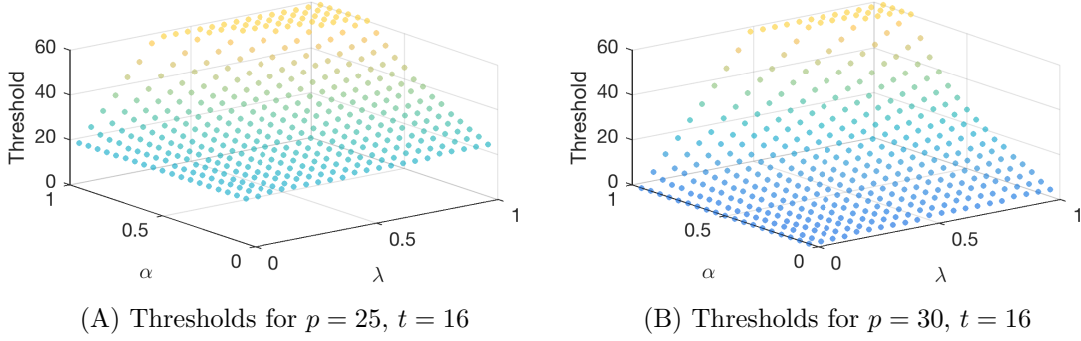


Figure 12: Behavior of Thresholds for Fixed p

6.3 Selecting β using Approximation

One way to choose β would be to compute and simulate a large batch of risk-averse policies corresponding to different values of β , similar to what is shown in Figure 12 and the references discussed in Section 1. As discussed in Section 5, this heuristic tuning method suffers from two major drawbacks: (1) it can be very computationally intense, due to the large number of β 's that needs to be explored, and (2) the final choice of β necessarily comes from the pre-selected batch. For instance, a MATLAB implementation on two 14 core, 3.3 GHz Intel Haswell CPUs requires 2.5 minutes for each instance of β (i.e., 13 MDPs with different horizons T). Solving these MDPs for 381 combinations of λ and α thus requires nearly 16 hours of CPU time (or without parallelization, *days* of CPU time).

We now illustrate the alternative approach that was outlined in Section 5, wherein a small sample of β 's are chosen as “observations” that are fed into a structured regression model to estimate the practical reward and risk over the entire set $\bar{\mathcal{B}}$. We assume that the practical reward is given by (5), the negative of the total expected cost. Let $\delta = 0.3$ and suppose the practical risk is given by

$$\rho_{\text{base}}[f_{\tau}^{-}(\mathbf{S}_{\tau}^{\pi})] = \mathbf{E}[\mathbf{1}_{\{R_{\tau}/R_{\max} \leq (1-\delta)\}}] = \mathbf{P}(R_{\tau}/R_{\max} \leq (1-\delta)),$$

the probability that the vehicle is less than 70% charged.⁵ Let $\varepsilon \in [0, 1]$ be the “trade-off parameter.” We consider the case where the decision maker aims to solve the practical problem (6) by appealing to the framework described in Figure 2 and searching over policies π that optimize *some instance* of the risk-averse MDP (similar to what is done in Maceira et al. [2015]). We proceed with the three steps outlined in Section 5.

Step 1. We let $\{\beta^1, \beta^2, \dots, \beta^n\}$ be all combinations of $\lambda \in \{0, 0.1, 0.2, \dots, 1\}$ and $\alpha \in \{0.05, 0.15, \dots, 0.95\}$. The corresponding MDPs are solved and the policies are simulated $N = 100,000$ times to obtain estimates of the practical reward and risk. Note that we are

⁵We choose 70% because a charge of around 80% is usually considered an adequate charge. For example, Tesla states that “charging above 80% isn’t typically necessary” at <http://www.tesla.com/supercharger>.

using a coarser grid than before to save computation time, allowing the regression procedure to fill in the blanks.

Step 2. We elect to use a polynomial representation for $\hat{r}_{\text{base}}(\beta)$ and $\hat{\rho}_{\text{base}}(\beta)$, so that

$$\hat{r}_{\text{base}}(\beta) = \mathbf{w}_{\Gamma}^{\top} \phi(\lambda, \alpha) \quad \text{and} \quad \hat{\rho}_{\text{base}}(\beta) = \mathbf{w}_{\Psi}^{\top} \phi(\lambda, \alpha),$$

where \mathbf{w}_{Γ} and \mathbf{w}_{Ψ} are the weight vectors and $\phi(\lambda, \alpha)$ is the vector of all degree 10 monomials (66 terms).

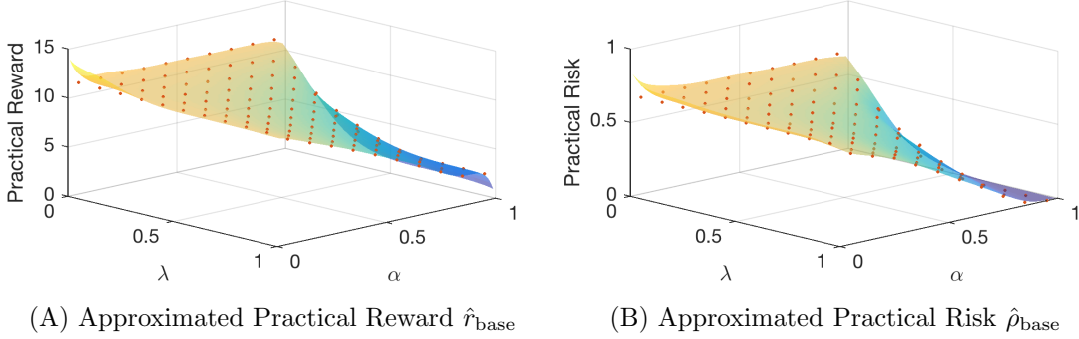


Figure 13: Practical Reward and Risk Metrics using Regression

Using the SPOT toolbox [Megretski, 2013] and the MOSEK solver, we carried out an l_1 regression procedure to compute the coefficients $\mathbf{w}_{\Gamma}, \mathbf{w}_{\Psi}$ — as mentioned in Section 5, sum-of-squares constraints [Ahmadi and Majumdar, 2014, 2015] are used to compute \mathbf{w}_{Ψ} to enforce the compatibility result of Theorem 5. The regression procedure requires roughly four seconds of CPU time (2 seconds for the reward and 2 seconds for the risk).

Figure 13 shows our estimated practical metric surface over $\beta \in \bar{\mathcal{B}}$ along with the metrics associated with the sample set of MDPs as red dots. In particular, we point out that the policy obtained by solving the *risk-neutral* (RN) MDP, i.e., when $\lambda = 0$, generates a profit of \$11.30, but leaves the EV inadequately charged 65.5% of the time.⁶ This level of risk would not be acceptable in practical settings.

Step 3. In the final step of the procedure, we use the approximations $\mathbf{w}_{\Gamma}, \mathbf{w}_{\Psi}$ to select a choice of β . Specifically, we solve the problem

$$\text{maximize}_{\beta \in [0,1] \times (0,1)} \hat{r}_{\text{base}}(\beta) \quad \text{subject to} \quad \hat{\rho}_{\text{base}}(\beta) \leq \varepsilon,$$

where the maximizer is denoted $\hat{\beta}(\varepsilon) = (\hat{\lambda}(\varepsilon), \hat{\alpha}(\varepsilon))$. This can be thought of as a “recommended degree of risk-aversion” for the particular level of ε . The policy selected for eventual implementation is $\pi_{\tau}^{\hat{\beta}(\varepsilon)}$ — the optimal policy with respect to the risk-averse MDP specified using $\hat{\beta}(\varepsilon)$. Of course, the solution $\hat{\beta}(\varepsilon)$ is generally not one of the sampled points,

⁶One may point out that surely we could change the inconvenience compensation scheme to mitigate this behavior. However, the idea of risk-averse optimization is to take the “physical characteristics” of the problem *as given* (i.e., fix the current compensation scheme) and produce risk-averse policies.

i.e., $\hat{\beta}(\varepsilon) \notin \{\beta^1, \beta^2, \dots, \beta^n\}$, so the final step before implementation is to solve the MDPs associated with $\hat{\beta}(\varepsilon)$. The recommended $\hat{\beta}(\varepsilon)$ and the practical reward and risk of the associated optimal policies for ten reasonable values for ε , ranging from 0.01 to 0.10, are shown in Table 1. The table also provides a comparison to the default policy (charge continuously) and the optimal risk-neutral (RN) solution, which we have given in the first and last rows, respectively.

| Trade-off ε | $\hat{\lambda}(\varepsilon)$ | $\hat{\alpha}(\varepsilon)$ | Reward (% of RN) | Risk (% of RN) |
|-------------------------|------------------------------|-----------------------------|------------------|----------------|
| Default | - | - | \$1.90 (16.8%) | 0.000 (0.0%) |
| 0.01 | 0.937 | 0.914 | \$2.74 (24.3%) | 0.002 (0.3%) |
| 0.02 | 0.876 | 0.901 | \$2.85 (25.3%) | 0.009 (1.3%) |
| 0.03 | 0.827 | 0.896 | \$2.97 (26.3%) | 0.019 (2.9%) |
| 0.04 | 0.793 | 0.891 | \$3.11 (27.5%) | 0.033 (5.0%) |
| 0.05 | 0.761 | 0.888 | \$3.25 (28.8%) | 0.041 (6.2%) |
| 0.06 | 0.747 | 0.881 | \$3.39 (30.0%) | 0.054 (8.2%) |
| 0.07 | 0.715 | 0.881 | \$3.52 (31.2%) | 0.068 (10.4%) |
| 0.08 | 0.681 | 0.883 | \$3.66 (32.4%) | 0.072 (10.9%) |
| 0.09 | 0.669 | 0.878 | \$3.80 (33.7%) | 0.090 (13.7%) |
| 0.10 | 0.61 | 0.89 | \$3.94 (34.9%) | 0.095 (14.5%) |
| RN | - | - | \$11.30 (100.0%) | 0.655 (100.0%) |

Table 1: Practical Reward and Risk of the Recommended Risk-Aversion Levels

Next, we show in Figure 14 the path of $\hat{\beta}(\varepsilon) = (\hat{\lambda}(\varepsilon), \hat{\alpha}(\varepsilon))$ as ε varies from 0.05 up to 0.6 (since we know that the RN solution incurs a practical risk of 0.655). Note that this figure is purely for methodological interest, as it is unlikely that ε values greater than 0.10 would be useful in practice. Although approximations from the regression are used to create the path of β , Figure 14 suggests that the relationship between (6) and the risk-averse MDP policies is complex and that there may not a structure that is simple to discern. Additional research in this area is needed.

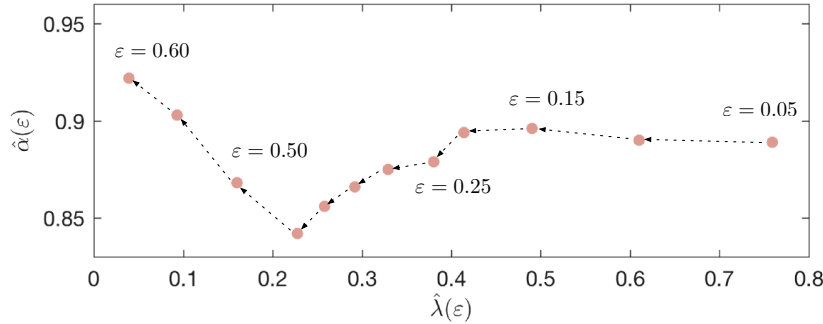


Figure 14: The Path of $\hat{\beta}$

7 Insights and Discussion

Our problem focuses on dynamic EV charging in the presence of a reservation system. However, “dynamic EV charging” may occur under various other related settings and system designs, and we believe that the results from our study apply there as well. Examples of these related problems are given below, along with a commentary on how our model can be extended to these settings.

- In a first-come, first-served EV charging station without reservations, the control system may attempt to “guess” at the return time of the customer and attempt to delay charging to reduce costs. In this model, the distribution of τ (the customer arrival time) is incorporated into the MDP, while in our current model, the distribution of τ defines multiple instances of the MDP.
- When a charging station has multiple ports that share the same power source (see, e.g., ChargePoint Express Plus and ChargePoint Power Cube), it is necessary to decrease the power output when additional vehicles join, creating the possibility that vehicles are not charged up to the customers’ expectations. Our model can be extended to this case by specifying a pre-determined rule for how power capacity should be split between vehicles; for example, in the interest of fairness, the rule could be that the power is always split evenly. Our model could then be applied independently at each port with the addition that the charging capacity x_{\max} is now a random variable (that decreases whenever additional vehicles are plugged into the station and increases when vehicles leave).
- Vehicle-to-grid systems take advantage of the storage capacity of idle electric vehicles, allowing them to provide ancillary services to the grid in times of need [Rotering and Ilic, 2011]. For our model to handle this case, we would allow the charging decision x_t to take both positive and negative values (negative values represent the delivery of energy to the grid).

A common thread of the above is that the customer is potentially disappointed when the charge level is inadequate at the time of use. Compensation for the inconvenience, in whatever form, monetary or otherwise, should then be provided. In this paper, we show an example where, under a reasonable compensation scheme, the optimal risk-neutral policy to the dynamic EV charging problem leaves the vehicle inadequately charged nearly *two-thirds of the time*. This suggests that for the broader class of dynamic EV charging problems, simply optimizing with respect to the expected total cost under any given inconvenience compensation scheme may produce undesirable policies.

One may point out that a trade-off between charging costs and “risk” (in the form of the compensation) is already being made in the *risk-neutral problem*. This is true to an extent; the issue is that trade-off induced by these competing costs is, in some sense, misaligned, as

the firm cannot implement a policy that provides an inadequate charge 65.5% of the time. Thus, the question is: if we take the physical characteristics of the problem as fixed (i.e., we do not re-write the compensation scheme), then how can we produce a policy that is more risk-averse? The introduction of risk measures provides an additional lever to “tip the scales” when the cost function itself provides an incorrect trade-off (utility functions are another approach). By implementing our risk-averse MDP/practical metrics approach on our reservation-based EV charging model, we are able to discover policies with much more reasonable trade-offs (see Table 1).

Most EV charging stations today use the default continuous charging policy, so it is perhaps more informative for our results to be viewed from the point of view of this standard practice. We see from Table 1 that if the manager of the charging station is willing to incur a small risk of 0.002, then a dynamic charging policy can increase profits by 44% (from \$1.90 to \$2.74). This suggests that EV charging stations can be operated at a significantly higher level of profitability if dynamic charging is adopted.

8 Conclusion

In this paper, we study EV charging in a setting where one attempts to exploit volatility in the spot market by charging only when the prices are low. At the same time, the risk of unsatisfied customers who may return to undercharged vehicles must be controlled. The problem is specified as a risk-averse MDP formulated using a dynamic (nested) risk measure, but we point out that the objective function has no practical interpretation. Our main methodological contribution lies in formalizing a procedure wherein risk-averse MDPs are used as a computational tool to produce policies that are evaluated using “practical” reward and reward metrics (termed the *base model*). We define a notion of *risk compatibility* and prove that under certain conditions, additional risk-aversion in the dynamic risk measure sense leads to additional risk-aversion in the practical sense as well. Next, we propose an approximation procedure that allows a decision-maker to easily search within a set of risk-averse MDPs for a “practical policy” using a regression model on a few solved instances. Finally, we provide numerical results based on real data that support the effectiveness of the approximation procedure. The numerical experiments suggest that dynamic EV charging can significantly improve the profitability of EV charging stations when a small amount of risk is incurred.

A Proofs

Proposition 1 (Convexity). *Let $s = (r, p) \in \mathcal{S}$ be the state variable. The following statements hold for every t and β : (i) $x p + \tilde{V}_{t,T}(r+x, p | \beta)$ is convex in x on $\mathcal{X}(r)$, (ii) $\tilde{V}_{t,T}(s | \beta)$ is convex in the resource r , and (iii) $V_{t,T}(s | \beta)$ is convex in the resource r .*

Proof. First note that $V_{t,T}(s | \beta)$ is trivially convex in r . We let the induction hypothesis be that $V_{t+1,T}(\cdot | \beta)(s)$ is strictly convex in r , from which we can show (i), (ii), and (iii) for t . For a feasible decision $x \in \mathcal{X}(r)$, let

$$Q_{t,T}(s, x | \beta) = x p - c_f + \tilde{V}_{t,T}(r+x, p | \beta)$$

and $\mathcal{Q} = \{(r, x) : x \in \mathcal{X}(r), r \in \mathcal{R}\}$. From the convexity of the mapping $(r, x) \mapsto V_{t+1,T}(r+x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)$ on \mathcal{Q} for each realization of ψ_{t+1} and the convexity (in the space of random variables) of ρ_{β_t} , we obtain that

$$(r, x) \mapsto \tilde{V}_{t,T}(r+x, p | \beta) = \rho_{\beta_t} [V_{t+1,T}(r+x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)]$$

is convex on \mathcal{Q} . This implies property (ii) and it follows that $Q_{t,T}(s, x | \beta)$ is convex in (r, x) on \mathcal{Q} , which verifies (i).

Since we can write $V_{t,T}(s | \beta) = \min_{x \in \mathcal{X}(r)} Q_{t,T}(s, x | \beta)$ and $\mathcal{X}(r)$ is nonempty, the well-known property that convexity is preserved under minimization can be applied; by [Heyman and Sobel, 2003, Proposition B-4], $V_{t,T}(s | \beta)$ is convex in r . This proves property (iii) and we have completed the inductive step and also verified (i) and (ii) along the way. \square

Proposition 2 (Monotonicity). *For every t and resource level $r \in \mathcal{R} \setminus \{R_{\max}\}$, the optimal value function $V_{t,T}(r, p | \beta)$ is increasing in the spot price p .*

Proof. We first focus on monotonicity in p and proceed via induction. For each resource state $r \in \mathcal{R} \setminus \{R_{\max}\}$, the function $V_{T,T}(s | \beta)$ is increasing in the spot price p by Assumption 1. This completes the base case. We assume the same is true for $V_{t+1,T}(r, p | \beta)$ and complete the inductive step by proving that the property holds for $V_{t,T}(r, p | \beta)$. To do so, we consider

$$\tilde{V}_{t,T}(r+x, p | \beta) = (1 - \lambda_t) \mathbf{E}[V_{t+1,T}(r+x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)] + \lambda_t c_{t,T}(r+x, p | \beta).$$

Since λ can take values of 0 or 1, we analyze both terms. By the induction hypothesis, the terms within the expectation are increasing in p for every realization of ψ_{t+1} , so the first term is increasing in p .

Now, we consider $c_{t,T}(r+x, p | \beta)$. Let $p' > p$ so that by the induction hypothesis, we know there exists a deterministic $\epsilon' > 0$ such that

$$V_{t+1,T}(r+x, p e^{-\kappa_Y} + \psi_{t+1} | \beta) + \epsilon' \leq V_{t+1,T}(r+x, p' e^{-\kappa_Y} + \psi_{t+1} | \beta) \quad a.s.$$

Hence, by the translation invariance and monotonicity axioms of coherent risk measures Artzner et al. [1999], we have $c_{t,T}(r+x, p | \beta) + \epsilon' \leq c_{t,T}(r+x, p' | \beta)$. We have shown that

$\tilde{V}_{t,T}(r+x, p | \beta)$ is increasing. Thus, $xp - c_f + \tilde{V}_{t,T}(r+x, p | \beta)$ is increasing in p for every $x \in \mathcal{X}(r)$, implying that $V_{t,T}(\cdot | \beta)$ is too. \square

Theorem 2 (Optimal Policy). *For each time $t < T$ and spot price p , there exists a threshold resource level $r_{t,T}(p | \beta)$ such that the optimal policy takes the following form:*

$$x_{t,T}(r, p | \beta) = \min\{r_{t,T}(p | \beta) - r, x_{\max}\} \mathbf{1}_{\{r \leq r_{t,T}(p | \beta)\}},$$

where $r_{t,T}(p | \beta)$ is the smallest value of \tilde{r} that minimizes the quantity $\tilde{r}p + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$.

Proof. To obtain an equivalent formulation of (11) in terms of the post-decision resource state, we make a substitution $\tilde{r} = r + x$ and $x = \tilde{r} - r$. Define $\mathcal{R}(r) = \{\tilde{r} : r \leq \tilde{r} \leq \min\{r + x_{\max}, R_{\max}\}\}$ as the set of possible post-decision resource levels. We have

$$x_{t,T}(r, p | \beta) \in \left(\arg \min_{\tilde{r} \in \mathcal{R}(r)} \tilde{r}p + \tilde{V}_{t,T}(\tilde{r}, p | \beta) \right) - r. \quad (14)$$

Note that the objective is convex in \tilde{r} by Proposition 1. Since $\mathcal{R}(r) \subseteq \mathcal{R}$, consider the solution $r_{t,T}(p | \beta)$ to the relaxed optimization problem

$$r_{t,T}(p | \beta) \in \arg \min_{\tilde{r} \in \mathcal{R}} \tilde{r}p + \tilde{V}_{t,T}(\tilde{r}, p | \beta). \quad (15)$$

Clearly, if $r_{t,T}(p | \beta) \in \mathcal{R}(r)$, then it is an optimal solution to the optimization problem within (14) and the optimal decision is $r_{t,T}(p | \beta) - r$. There are two ways in which $r_{t,T}(p | \beta) \notin \mathcal{R}(r)$ can happen. First, if $r_{t,T}(p | \beta) < r$, then by convexity, $\tilde{r}p + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$ is nondecreasing on $\mathcal{R}(r)$ and an optimal solution is $\tilde{r} = r$, implying an optimal decision of 0. The second case is when $r + x_{\max} < r_{t,T}(p | \beta) \leq R_{\max}$. Again, by convexity, it must be true that $\tilde{r}p + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$ is nonincreasing on $\mathcal{R}(r)$ and an optimal solution is $\tilde{r} = r + x_{\max}$. This necessarily gives us an optimal decision of x_{\max} . Finally, combining these three cases allows us to conclude the structure of the optimal policy stated in the theorem. \square

Lemma 1 (Lipschitz in Spot Price). *If Assumption 1 holds, then for any t and r , the value function $V_{t,T}(r, \cdot | \beta)$ is Lipschitz continuous in the price dimension.*

Proof. First, the terminal cost function

$$V_{t,T}(r, p | \beta) = [1 + \gamma_h h_t(r) + \rho_{\beta_T}[\gamma_Y(p e^{-\kappa_Y} + \psi_{T+1,Y})]] h_t(r) p_{\text{ref}}$$

is clearly Lipschitz in p for each r by Assumption 1 and the inequality $|\rho_{\beta_t}(A) - \rho_{\beta_t}(B)| \leq \rho_{\beta_t}(|A - B|)$. For each fixed r , we denote the Lipschitz parameter by $L_{V_{t,T} | \beta}(r)$ and then let $L_{V_{t,T} | \beta} = \max_{r \in \mathcal{R}} L_{V_{t,T} | \beta}(r)$. We now proceed by backward induction on t . Using the inequality $|\min f - \min g| \leq \max |f - g|$, we have

$$|V_{t,T}(r, p | \beta) - V_{t,T}(r, p' | \beta)| \leq \max_{x \in \mathcal{X}(r)} |\tilde{V}_{t,T}(r+x, p | \beta) - \tilde{V}_{t,T}(r+x, p' | \beta)|. \quad (16)$$

Using the definition of the post-decision value function and (4), we have

$$\tilde{V}_{t,T}(r+x, p | \beta) = \rho_{\beta_t} [V_{t+1,T}(r+x, p e^{-\kappa_Y} + \psi_{t+1} | \beta)].$$

Again using the inequality $|\rho_{\beta_t}(A) - \rho_{\beta_t}(B)| \leq \rho_{\beta_t}(|A - B|)$ along with (16), we conclude that

$$|V_{t,T}(r+x, p | \beta) - V_{t,T}(r+x, p' | \beta)| \leq e^{-\kappa_Y} L_{V_{t+1,T} | \beta} |p - p'|$$

for any $r+x$. Thus, we can set $L_{V_{t,T} | \beta} = e^{-\kappa_Y} L_{V_{t+1,T} | \beta}$ to complete the proof. \square

Lemma 2 (Differentiability of Value at Risk). *For every resource state $r \in \mathcal{R} \setminus \{R_{\max}\}$, the value at risk function $v_{t,T}(r, p | \beta) = \text{VaR}_{\alpha_t}[V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)]$ is differentiable for almost every spot price p .*

Proof. We apply the simple property that value at risk is invariant under a monotone transformation in the following sense. Consider a random variable X and an increasing function l . Then, since $\mathbf{P}[X \leq \text{VaR}_{\alpha_t}(X)] = \mathbf{P}[l(X) \leq l(\text{VaR}_{\alpha_t}(X))] = \alpha$, it follows that $\text{VaR}_{\alpha_t}(l(X)) = l(\text{VaR}_{\alpha_t}(X))$. By the monotonicity property of Proposition 2, we have

$$v_{t,T}(r, p | \beta) = \text{VaR}_{\alpha_t}[V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)] = V_{t+1,T}(r, p e^{-\kappa_Y} + \text{VaR}_{\alpha_t}(\psi_{t+1}) | \beta)$$

We conclude by noting that the Lipschitz property of Lemma 1 implies $V_{t+1,T}(r, \cdot | \beta)$ is differentiable almost everywhere. \square

Lemma 3 (Differentiating under the Risk Measure). *Suppose the value function $V_{t+1,T}(\cdot | \beta)$ is differentiable at $r = r_0 \in \mathcal{R} \setminus \{R_{\max}\}$. Then, we may differentiate under the risk measure in the following sense:*

$$\begin{aligned} \partial_r \tilde{V}_{t,T}(r, p | \beta) |_{r=r_0} &= (1 - \lambda) \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)] |_{r=r_0} \\ &\quad + \lambda \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) | \psi_{t+1} \geq \text{VaR}_{\alpha_t}(\psi_{t+1})] |_{r=r_0}. \end{aligned}$$

Proof. The first part of the equation is simply an interchange of differentiation and expectation:

$$\partial_r \mathbf{E}[V_{t+1}^\beta(r, p e^{-\kappa_Y} + \psi_{t+1})] = \mathbf{E}[\partial_r V_{t+1}^\beta(r, p e^{-\kappa_Y} + \psi_{t+1})].$$

By Proposition 1, $V_{t+1}^\beta(r, p e^{-\kappa_Y} + \psi_{t+1})$ is Lipschitz on $[0, R_{\max}]$ and thus the interchange follows by the dominated convergence theorem. To show the second part, we first apply [Hong and Liu, 2009, Theorem 3.1] to obtain

$$\begin{aligned} &\partial_r \text{CVaR}_{\alpha_t}[V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta)] |_{r=r_0} \\ &= \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) | V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) \geq v_{t,\alpha}(r, p | \beta)] |_{r=r_0}. \end{aligned} \tag{17}$$

Invoking this theorem requires verifying [Hong and Liu, 2009, Assumptions 1-3]. [Hong and Liu, 2009, Assumptions 1-2] are verified immediately by Proposition 1 and Lemma 2. [Hong and Liu, 2009, Assumption 3] states that $\{V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) = v_{t,\alpha}(r, p | \beta)\}$ is a zero probability event, which is clear by Proposition 2. To finish the proof, we note that the conditioning event $\{V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) \geq v_{t,\alpha}(r, p | \beta)\}$ is equivalent to

$$\{V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1} | \beta) \geq V_{t+1,T}(r, p e^{-\kappa_Y} + \text{VaR}_{\alpha_t}(\psi_{t+1}) | \beta)\}$$

by monotonicity in the spot price (Proposition 2). Inverting both sides gives the desired characterization $\{\psi_{t+1} \geq \text{VaR}_{\alpha_t}(\psi_{t+1})\}$. \square

Theorem 3 (Nonincreasing Threshold in Spot Price). *Under Assumption 1, the threshold $r_{t,T}(p|\beta)$ is nonincreasing in the spot price p for any degree of dynamic risk-aversion β .*

Proof. The proof is by backward induction on t . The *induction hypothesis* is:

- The quantity $p + \partial_r \tilde{V}_{t,T}(r, p|\beta)$ is nondecreasing.

Due to the characterization of $r_{t,T}(p|\beta)$ as the solution to the minimization problem $\tilde{r}p + \tilde{V}_{t,T}(\tilde{r}, p|\beta)$ over $\tilde{r} \in \mathcal{R}$, this induction hypothesis implies that $r_{t,T}(p|\beta)$ is nonincreasing in p .

Some Remarks. Before continuing with the proof, we make some remarks regarding differentiability in r and p . The value function $V_{t+1,T}(r, \cdot|\beta)$ is *piecewise continuously differentiable*⁷ in the resource state r . This can be shown via an induction argument combined with applications of Lemma 3, the envelope theorem, and the fact that $x_{t,T}(r, p|\beta)$ is piecewise linear. Thus, our argument centers on resource states in the interval $(0, R_{\max})$ at which differentiability holds. We then extend the result of the theorem to directional derivatives at the nondifferentiable endpoints/breakpoints by invoking piecewise continuous differentiability and taking limits for each “piece.”

Base Case. We first consider the base case, $t+1 = T$, i.e., the case of the terminal value function $V_{t,T}(s|\beta)$. By Assumption 1, the derivative

$$\begin{aligned} \partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1}|\beta) \\ = -p_{\text{ref}}[1 + 2\gamma_h h(r) + \rho_{\beta_T}[\gamma_Y((p e^{-\kappa_Y} + \psi_{t+1}) e^{-\kappa_Y} + \psi_{t+1,Y})]] \end{aligned}$$

is a decreasing function in p , and invoking the inequality $|\rho_{\beta_t}(A) - \rho_{\beta_t}(B)| \leq \rho_{\beta_t}(|A - B|)$, we have that for any p and q ,

$$|\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1}|\beta) - \partial_r V_{t+1,T}(r, q e^{-\kappa_Y} + \psi_{t+1}|\beta)| \leq p_{\text{ref}} L_{\gamma_Y} e^{-2\kappa_Y} |p - q|.$$

Since the Lipschitz constant satisfies $L_{\gamma_Y} \leq p_{\text{ref}}^{-1} e^{2\kappa_Y}$ (by Assumption 1), the right-hand-side of the inequality above can be bounded by $|p - q|$. By Lemma 3, we have

$$\begin{aligned} \partial_r \tilde{V}_{t,T}(r, p|\beta) &= (1 - \lambda_t) \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1}|\beta)] \\ &\quad + \lambda_t \mathbf{E}[\partial_r V_{t+1,T}(r, p e^{-\kappa_Y} + \psi_{t+1}|\beta) | \psi_{t+1} \geq \text{VaR}_{\alpha_t}(\psi_{t+1})], \end{aligned}$$

⁷The notion of *piecewise continuously differentiability* for some function $f : [a, b] \rightarrow \mathbb{R}$ states that there exists a_i with $a = a_0 < a_1 < \dots < a_n = b$ such that f is continuously differentiable when restricted to $[a_i, a_{i+1}]$ for each $i < n$. More precisely, for a fixed i , $f'(x)$ at $x \in (a_i, a_{i+1})$ is the derivative, while $f'(a_i)$ is the right derivative, and $f'(a_{i+1})$ is the left derivative and f' is continuous on $[a_i, a_{i+1}]$. By Theorem 2, the optimal policy $x_{t,T}(r, p|\beta)$ is piecewise linear and hence, piecewise continuously differentiable in r .

which satisfies $|\partial_r \tilde{V}_{t,T}(r, p | \beta) - \partial_r \tilde{V}_{t,T}(r, q | \beta)| \leq |p - q|$. Therefore, it follows that the quantity $p + \partial_r \tilde{V}_{t,T}(r, p | \beta)$ is nondecreasing in p .

Inductive Step (Part 1). Let us now consider the value function at time t evaluated at the spot price $P = p e^{-\kappa_Y} + \psi_t$ (i.e., from the perspective of time $t - 1$). Recall that the optimal policy at time t , by Theorem 2, is to charge up to (or as close as possible to) the threshold $r_{t,T}(p | \beta)$. We define the price sets $\mathcal{P}_{\text{low}}(r) = \{P : r + x_{\text{max}} < r_{t,T}(p | \beta)\}$, $\mathcal{P}_{\text{med}}(r) = \{P : r \leq r_{t,T}(p | \beta) \leq r + x_{\text{max}}\}$, and $\mathcal{P}_{\text{high}}(r) = \{P : r_{t,T}(p | \beta) < r\}$, where it is possible for some of these sets to be empty.

Due to $r_{t,T}(p | \beta)$ being nonincreasing in p , these price sets are disjoint and monotone in the sense that any element of $\mathcal{P}_{\text{low}}(r)$ is no greater than any element of $\mathcal{P}_{\text{med}}(r)$, and any element of $\mathcal{P}_{\text{med}}(r)$ is no greater than any element of $\mathcal{P}_{\text{high}}(r)$. See Figure 15 for an illustration.

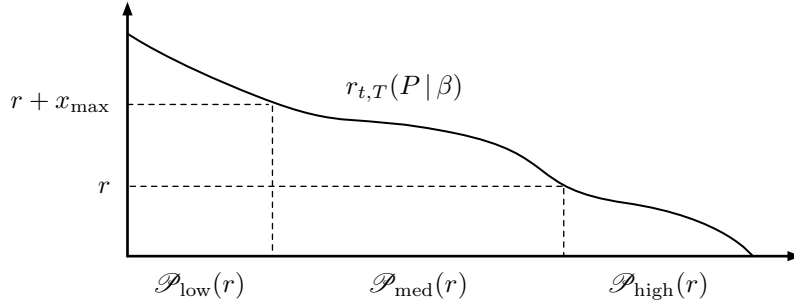


Figure 15: Illustration of the Price Sets

The value function can be written using the optimal policy and the post-decision value function:

$$V_{t,T}(r, p | \beta) = x_{t,T}(r, p | \beta) \cdot P - c_f + \tilde{V}_{t,T}^\beta(r + x_{t,T}(r, p | \beta), P).$$

Doing casework on each of the spot price intervals, we have

$$V_{t,T}(r, p | \beta) = \begin{cases} x_{\text{max}} \cdot P - c_f + \tilde{V}_{t,T}(r + x_{\text{max}}, P | \beta) & \text{if } P \in \mathcal{P}_{\text{low}}(r), \\ (r_{t,T}(p | \beta) - r) \cdot P - c_f + \tilde{V}_{t,T}(r_{t,T}(p | \beta), P | \beta) & \text{if } P \in \mathcal{P}_{\text{med}}(r), \\ -c_f + \tilde{V}_{t,T}(r, P | \beta) & \text{if } P \in \mathcal{P}_{\text{high}}(r). \end{cases}$$

Differentiating in r and adding P to both sides, we have

$$P + \partial_r V_{t,T}(r, p | \beta) = \begin{cases} P + \partial_r \tilde{V}_{t,T}(r + x_{\text{max}}, P | \beta) & \text{if } P \in \mathcal{P}_{\text{low}}(r), \\ 0 & \text{if } P \in \mathcal{P}_{\text{med}}(r), \\ P + \partial_r \tilde{V}_{t,T}(r, p | \beta) & \text{if } P \in \mathcal{P}_{\text{high}}(r). \end{cases}$$

Recall the characterization of $r_{t,T}(p | \beta)$ minimizes the quantity $\tilde{r} p + \tilde{V}_{t,T}(\tilde{r}, p | \beta)$. For $P \in \mathcal{P}_{\text{low}}(r)$, we can infer by convexity and the property $r + x_{\text{max}} < r_{t,T}(p | \beta)$ that $P + \partial_r \tilde{V}_{t,T}(r + x_{\text{max}}, P | \beta) \leq 0$. Similar reasoning yields $P + \partial_r \tilde{V}_{t,T}(r, p | \beta) \geq 0$ for $P \in$

$\mathcal{P}_{\text{high}}(r)$. Hence, $P + \partial_r V_{t,T}(r, p | \beta)$ is nondecreasing in P .

Inductive Step (Part 2). Given this, we are now able to more easily analyze the main quantity of interest, $p + \partial_r \tilde{V}_{T-1,T}(r, p | \beta)$. By Lemma 3, we have

$$\begin{aligned} p + \partial_r \tilde{V}_{T-1,T}(r, p | \beta) &= (1 - \lambda_t) [p + \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta)]] \\ &\quad + \lambda_t [p + \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta) | \psi_t \geq \text{VaR}_{\alpha_t}(\psi_t)]]. \end{aligned} \quad (18)$$

The first term in brackets can be written as

$$\begin{aligned} p + \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta)] &= p + \mathbf{E}[P + \partial_r V_{t,T}(r, p | \beta) - P] \\ &= (p - p e^{-\kappa_Y}) - \mathbf{E}(\psi_t) + \mathbf{E}[P + \partial_r V_{t,T}(r, p | \beta)], \end{aligned}$$

which is clearly nondecreasing in p given the property we proved in Part 1. The second term is the same, except with conditional expectations that do not depend on p . \square

Lemma 4 (Decreasing Marginal Value in Price). *Under Assumption 1, the pre-decision and post-decision marginal value of energy, $\partial_r V_{t,T}(r, p | \beta)$ and $\partial_r \tilde{V}_{t,T}(r, p | \beta)$, are nonincreasing in the spot price p for any fixed risk-aversion β .*

Proof. We proceed via induction. The base case, that $\partial_r \tilde{V}_{T-1,T}(r, p | \beta)$, is nonincreasing in the spot price can be easily verified by essentially repeating the proof for the base case of Theorem 3. For the inductive step, we assume the property holds at t and verify it for $t - 1$. Let $P = p e^{-\kappa_Y} + \psi_t$. Following the definitions from the proof of Theorem 3, we have

$$\partial_r V_{t,T}(r, p | \beta) = \begin{cases} \partial_r \tilde{V}_{t,T}(r + x_{\max}, P | \beta) & \text{if } P \in \mathcal{P}_{\text{low}}(r), \\ -P & \text{if } P \in \mathcal{P}_{\text{med}}(r), \\ \partial_r \tilde{V}_{t,T}(r, p | \beta) & \text{if } P \in \mathcal{P}_{\text{high}}(r). \end{cases}$$

Although this function is clearly piecewise nonincreasing from the induction hypothesis, we would like to verify that $\partial_r V_{t,T}(r, p | \beta)$ nonincreasing over all P . We first consider the term $\partial_r \tilde{V}_{t,T}(r + x_{\max}, P | \beta)$. When $P \in \mathcal{P}_{\text{low}}(r)$ and the optimal unconstrained charge level $r_{t,T}(p | \beta)$ is not reachable from r , we can apply the same reasoning as in the proof of Theorem 3 to deduce via convexity that $P + \partial_r \tilde{V}_{t,T}(r + x_{\max}, P | \beta) \leq 0$. When $P \in \mathcal{P}_{\text{med}}(r)$, we know that $r_{t,T}(p | \beta)$ is reachable from r and thus it follows that $P + \partial_r \tilde{V}_{t,T}(r + x_{\max}, P | \beta) \geq 0$. Rearranging these inequalities, we can see that

$$\partial_r V_{t,T}(r, p | \beta) = \min\{-P, \partial_r \tilde{V}_{t,T}(r + x_{\max}, P | \beta)\} \quad \text{if } P \in \mathcal{P}_{\text{low}}(r) \cup \mathcal{P}_{\text{med}}(r). \quad (19)$$

By the induction hypothesis, $\partial_r V_{t,T}(r, p | \beta)$ is the minimum of two nonincreasing functions on $P \in \mathcal{P}_{\text{low}}(r) \cup \mathcal{P}_{\text{med}}(r)$.

Now let us consider the term $\partial_r \tilde{V}_{t,T}(r, p | \beta)$. For $P \in \mathcal{P}_{\text{med}}(r)$, the optimal charge level is reachable and thus $r \leq r_{t,T}(p | \beta)$; by convexity $P + \partial_r \tilde{V}_{t,T}(r, p | \beta) \leq 0$. For $P \in \mathcal{P}_{\text{high}}(r)$, we have $r > r_{t,T}(p | \beta)$ and thus $P + \partial_r \tilde{V}_{t,T}(r, p | \beta) \geq 0$. Therefore,

$$\partial_r V_{t,T}(r, p | \beta) = \max\{-P, \partial_r \tilde{V}_{t,T}(r, p | \beta)\} \quad \text{if } P \in \mathcal{P}_{\text{med}}(r) \cup \mathcal{P}_{\text{high}}(r). \quad (20)$$

Combining (19) and (20), we see that $\partial_r V_{t,T}(r, p | \beta)$ is nonincreasing over all $P \in \mathbb{R}$. This proves the first part of the lemma. To conclude the second part, we invoke Lemma 3 to get

$$\begin{aligned} \partial_r \tilde{V}_{T-1,T}(r, p | \beta) &= (1 - \lambda_t) \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta)] \\ &\quad + \lambda_t \mathbf{E}[\partial_r V_{T,T}(r, p e^{-\kappa_Y} + \psi_T | \beta) | \psi_t \geq \text{VaR}_{\alpha_t}(\psi_t)]. \end{aligned}$$

The right-hand-side is clearly nonincreasing in p . \square

Theorem 4 (Compatibility in General Charging Regime). *Fix any reservation horizon T and the risk-aversion $\beta_{-0}^* = (\lambda_1^*, \alpha_1^*, \dots, \lambda_T^*, \alpha_T^*)$ for all periods except the first. Let*

$$\mathcal{B}(\beta_{-0}^*) = \{\beta = (\lambda_0, \alpha_0, \lambda_1^*, \alpha_1^*, \dots, \lambda_T^*, \alpha_T^*) : \lambda_0 \in [0, 1], \alpha_0 \in (0, 1)\}$$

be the set of risk-aversion parameters where (λ_t, α_t) matches $(\lambda_t^, \alpha_t^*)$ of β_{-0}^* while the risk-aversion in the first period is a “free parameter.” Taking the partial order \leq to be the componentwise inequality on \mathbb{R}^2 gives compatibility on the set $\mathcal{B}(\beta_{-0}^*)$ between $\rho_{0,T}(\cdot | \beta)$ and any resource-monotone f_T^- .*

Proof. By Lemma 3, we have the equation

$$\begin{aligned} \partial_r \tilde{V}_{0,T}(r, p | \beta) &= (1 - \lambda_0) \mathbf{E}[\partial_r V_{1,T}(r, p e^{-\kappa_Y} + \psi_1 | \beta)] \\ &\quad + \lambda_0 \mathbf{E}[\partial_r V_{1,T}(r, p e^{-\kappa_Y} + \psi_1 | \beta) | \psi_1 \geq \text{VaR}_{\alpha_0}(\psi_1)]. \end{aligned} \tag{21}$$

We note that by the dynamic programming property of Theorem 1, the value function $V_{1,T}^\beta$ at time $t = 1$ depends only on β_{-0}^* and not on the risk-aversion of the first time period $\beta_0 = (\lambda_0, \alpha_0)$. By the nonincreasing property of $\partial_r V_{1,T}^\beta(r, p)$ in p from Lemma 4, it follows that

$$\mathbf{E}[\partial_r V_{1,T}(r, p e^{-\kappa_Y} + \psi_1 | \beta) | \psi_1 \geq \text{VaR}_{\alpha_0}(\psi_1)] \leq \mathbf{E}[\partial_r V_{1,T}(r, p e^{-\kappa_Y} + \psi_1 | \beta)]$$

and $\alpha_0 \mapsto \mathbf{E}[\partial_r V_{1,T}(r, p e^{-\kappa_Y} + \psi_1 | \beta) | \psi_1 \geq \text{VaR}_{\alpha_0}(\psi_1)]$ is nonincreasing. Combining these two properties allows us to conclude from (21) that $\partial_r \tilde{V}_{0,T}(r, p | \beta)$ is nonincreasing in β over $\mathcal{B}(\beta_{-0}^*)$, which in turn implies that the charging threshold $r_{0,T}(p | \beta)$ is *nondecreasing*. The thresholds for $t > 0$ remain unchanged when β (again by the dynamic programming property) and thus, it is easy to conclude that $\mathbf{R}_T^\beta \leq \mathbf{R}_T^{\beta'}$ almost surely when $\beta \leq \beta'$. Risk compatibility follows by the resource-monotonicity property of f_T^- . \square

References

- B. Acciaio and I. Penner. Dynamic risk measures. In *Advanced Mathematical Methods for Finance*, pages 1–34. Springer Berlin Heidelberg, 2011.
- A. A. Ahmadi and A. Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *2014 IEEE Conference on Information Sciences and Systems (CISS)*, pages 1–5. IEEE, 2014.
- A. A. Ahmadi and A. Majumdar. Some applications of polynomial optimization in operations research and real-time decision making. *Optimization Letters*, pages 1–21, 2015.
- S. Ahmed, U. Çakmak, and A. Shapiro. Coherent risk measures in inventory problems. *European Journal of Operational Research*, 182(1):226–238, 2007.
- R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3: 5–40, 2001.
- E. Altman. *Constrained Markov decision processes*, volume 7. CRC Press, 1999.
- P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.
- S. Bashash, S. J. Moura, and H. K. Fathy. Charge trajectory optimization of plug-in hybrid electric vehicles for energy cost reduction and battery health enhancement. *American Control Conference (ACC), 2010*, pages 5824–5831, 2010.
- S. Bashash, S. J. Moura, J. C. Forman, and H. K. Fathy. Plug-in hybrid electric vehicle charge pattern optimization for energy cost and battery longevity. *Journal of Power Sources*, 196(1):541–549, 2011.
- Bay Area Council. Electrify your business: Moving forward with electric vehicles – A Bay area business guide. 2011.
- I. S. Bayram, G. Michailidis, M. Devetsikiotis, S. Bhattacharya, A. Chakraborty, and F. Granelli. Local energy storage sizing in plug-in hybrid electric vehicle charging stations under blocking probability constraints. *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 78–83, 2011.
- J. Belles-Sampera, M. Guillén, and M. Santolino. Beyond value-at-risk: GlueVaR distortion risk measures. *Risk Analysis*, 34(1):121–134, 2014.
- K. Boda and J. A. Filar. Time consistent dynamic risk measures. *Mathematical Methods of Operations Research*, 63(1):169–186, 2006.

- M. Bouakiz and M. J. Sobel. Inventory control with an exponential utility criterion. *Operations Research*, 40(3):603–608, 1992.
- N. Caperello, K. S. Kurani, and J. TyreeHageman. Do you mind if I plug-in my car? How etiquette shapes PEV drivers’ vehicle charging behavior. *Transportation Research Part A: Policy and Practice*, 54:155–163, 2013.
- Á. Cartea and M. G. Figueroa. Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4):313–335, 2005.
- O. Çavus and A. Ruszczyński. Computational methods for risk-averse undiscounted transient markov models. *Operations Research*, 62(2):401–417, 2014.
- D. Chang, D. Erstad, E. Lin, A. F. Rice, C. T. Goh, A. A. Tsao, and J. Snyder. Financial viability of non-residential electric vehicle charging stations. *Luskin Center for Innovation: Los Angeles, CA, USA.*, 2012.
- D. Chen, S. Benjaafar, and W. L. Cooper. On the Impact of Input Price Variability in Stochastic Inventory Systems. *Working paper, University of Minnesota*, 2015.
- X. Chen, M. Sim, D. Simchi-Levi, and P. Sun. Risk aversion in inventory management. *Operations Research*, 55(5):828–842, 2007.
- Y. F. Chen, M. Xu, and Z. G. Zhang. Technical note - A risk-averse newsvendor model under the CVaR criterion. *Operations Research*, 57(4):1040–1044, 2009.
- P. Cheridito and M. Stadje. Time-inconsistency of VaR and time-consistent alternatives. *Finance Research Letters*, 6(1):40–46, 2009.
- P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11(3):57–106, 2006.
- S. Choi, A. Ruszczyński, and Y. Zhao. A multiproduct risk-averse newsvendor with law-invariant coherent measures of risk. *Operations Research*, 59(2):346–364, 2011.
- K. Clement-Nyns, E. Haesen, and J. Driesen. The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. *IEEE Transactions on Power Systems*, 25(1):371–380, 2010.
- R. Collado, S. Meisel, and L. Priekule. Risk-averse stochastic path detection. *European Journal of Operational Research*, 260(1):195–211, 2017.
- H. Dette, N. Neumeyer, and K. F. Pilz. A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli*, 12(3):469–490, 2006.
- S. K. Devalkar, R. Anupindi, and A. Sinha. Dynamic risk management of commodity operations: Model and analysis. *Working Paper, Indian School of Business*, 2016.

- L. Eeckhoudt, C. Gollier, and H. Schlesinger. The risk-averse (and prudent) newsboy. *Management Science*, 41(5):786–794, 1995.
- A. Eydeland and K. Wolyniec. *Energy and Power Risk Management*. Wiley, Hoboken, New Jersey, 2003.
- M. Frittelli and E. R. Gianin. Dynamic convex risk measures. In *Risk measures for the 21st century*, volume 8, pages 227–248. 2004.
- L. Gan, U. Topcu, and S. Low. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems*, 28(2):940–951, 2013.
- J. Gotoh and Y. Takano. Newsvendor solutions via conditional value-at-risk minimization. *European Journal of Operational Research*, 179(1):80–96, 2007.
- D. P. Heyman and M. J. Sobel. *Stochastic Models in Operations Research: Stochastic Optimization*, volume 2. Courier Corporation, 2003.
- L. J. Hong and G. Liu. Simulating sensitivities of conditional value at risk. *Management Science*, 55(2):281–293, 2009.
- D. A. Iancu, M. Petrik, and D. Subramanian. Tight approximations of dynamic risk measures. *Mathematics of Operations Research*, 40(3):655–682, 2015.
- E. B. Iversen, J. M. Morales, and H. Madsen. Optimal charging of an electric vehicle using a Markov decision process. *Applied Energy*, 123:1–12, 2014.
- M. M. Karbasioun, I. Lambadaris, G. Shaikhet, and E. Kranakis. Optimal charging strategies for electrical vehicles under real time pricing. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 746–751, 2014.
- J. H. Kim and W. B. Powell. Optimal energy commitments with storage and intermittent supply. *Operations Research*, 59(6):1347–1360, 2011.
- I. Koutsopoulos, V. Hatzi, and L. Tassiulas. Optimal energy storage control policies for the smart power grid. *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 475–480, 2011.
- R. Li, Q. Wu, and S. S. Oren. Distribution locational marginal pricing for optimal electric vehicle charging management. *IEEE Transactions on Power Systems*, 29(1):203–211, 2014.
- Q. Lin, X. Chen, and J. Pena. A trade execution model under a composite dynamic coherent risk measure. *Operations Research Letters*, 43(1):52–58, 2015.

- M. E. P. Maceira, L. G. B. Marzano, D. D. J. Penna, A. L. Diniz, and T. C. Justino. Application of CVaR risk aversion approach in the expansion and operation planning and for setting the spot price in the Brazilian hydrothermal interconnected system. *International Journal of Electrical Power and Energy Systems*, 72:126–135, 2015.
- V. Martínez-De-Albéniz and D. Simchi-Levi. Mean-variance trade-offs in supply contracts. *Naval Research Logistics*, 53(7):603–616, 2006.
- A. Megretski. SPOT: Systems polynomial optimization tools, 2013.
- H. Mukerjee. Monotone nonparametric regression. *The Annals of Statistics*, 16(2):741–750, 1988.
- P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming, Series B*, 96(2):293–320, 2003.
- G. C. Pflug and A. Ruszczyński. Measuring risk for income streams. *Computational Optimization and Applications*, 32(1-2), 2005.
- A. B. Philpott and V. L. de Matos. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2):470–483, 2012.
- A. B. Philpott, V. L. de Matos, and E. Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4):957–970, 2013.
- M. L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. Wiley, New York, 2014.
- W. Qi, Z. Xu, Z. J. M. Shen, Z. Hu, and Y. Song. Hierarchical coordinated control of plug-in electric vehicles charging in multifamily dwellings. *IEEE Transactions on Smart Grid*, 5(3):1465–1474, 2014.
- F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200, 2004.
- J. Robinson, G. Brase, W. Griswold, C. Jackson, and L. Erickson. Business models for solar powered charging stations to develop infrastructure for electric vehicles. *Sustainability*, 6(10):7358–7387, 2014.
- R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.
- C. Roe, E. Farantatos, J. Meisel, A. P. Meliopoulos, and T. Overbye. Power system level impacts of PHEVs. *2009 42nd Hawaii International Conference on System Sciences*, pages 1–10, 2009.

- N. Rotering and M. Ilic. Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets. *IEEE Transactions on Power Systems*, 26(3):1021–1029, 2011.
- D. Roy. The discrete normal distribution. *Communications in Statistics – Theory and Methods*, 32(10):1871–1883, 2003.
- B. Rudloff, A. Street, and D. M. Valladão. Time consistency and risk averse dynamic decision models: definition, interpretation and practical consequences. *European Journal of Operational Research*, 234(3):743–750, 2014.
- A. Ruszczyński. Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming*, 125(2):235–261, 2010.
- A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, 2006.
- A. Schroeder and T. Traber. The economics of fast charging infrastructure for electric vehicles. *Energy Policy*, 43:136–144, 2012.
- N. Secomandi. Optimal commodity trading with a capacitated storage asset. *Management Science*, 56(3):449–467, 2010.
- A. Shapiro, W. Tekaya, J. P. da Costa, and M. P. Soares. Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224(2):375–391, 2013.
- E. Sortomme, M. M. Hindi, S. D. James MacPherson, and S. S. Venkata. Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses. *IEEE Transactions on Smart Grid*, 2(1):186–193, 2011.
- O. Sundstrom and C. Binding. Optimization methods to plan the charging of electric vehicle fleets. *Proceedings of the International Conference on Control Communication and Power Engineering*, pages 28–29, 2010.
- U.S. Department of Energy. Plug-in electric vehicle handbook for public charging station hosts. Technical report, 2012.
- M. D. Wytock, K. Adelberger, W. Wolf-Dietrich, and O. Irzak. Mediator device for smart electric vehicle charging. *US Patent Application No. 14/928,170*, 2016.
- J. Yang. Monotone trends in inventory-price control under time-consistent coherent risk measure. *Operations Research Letters*, 45(3):293–299, 2017.

- T. Zhang, W. Chen, Z. Han, and Z. Cao. Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price. *IEEE Transactions on Vehicular Technology*, 63(6):2600–2612, 2014.
- Y. H. Zhou, A. Scheller-Wolf, N. Secomandi, and S. Smith. Electricity trading and negative prices: Storage vs. disposal. *Management Science*, 2015.