Fall 12-16-2023

Essays in Robust Optimization with Applications to Finance and Renewable Energy

Hao Jiang
haojiang@smu.edu

ESSAYS IN ROBUST OPTIMIZATION

WITH APPLICATIONS TO FINANCE AND RENEWABLE ENERGY

Approved by:

Dr. Aurélie Thiele
Associate Professor

Dr. Miju Ahn
Assistant Professor

Dr. Digvijay Boob
Assistant Professor

Dr. Michael Hahsler
Clinical Associate Professor

Dr. Eojin Han
Assistant Professor

ESSAYS IN ROBUST OPTIMIZATION
WITH APPLICATIONS TO FINANCE AND RENEWABLE ENERGY

A Dissertation Presented to the Graduate Faculty of the

Lyle School of Eneineering

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Operations Research

by

Hao Jiang

B.S., Economics and Mathematics, University of Kansas
M.A., Economics and Applied Statistics, UC Santa Barbara

December 16, 2023

Acknowledgments

First, I would like to express my gratitude to my advisor, Dr. Aurélie Thiele, whose expertise and guidance have been the foundation of my research at Southern Methodist University. Her profound knowledge and constant encouragement have been invaluable in helping me to tackle the challenges of my work.

I am grateful to Dr. Miju Ahn, Dr. Digijay Boob, Dr. Michael Hahsler, and Dr. Eojin Han for their contributions to my research. Their insightful suggestions have greatly enriched my work. I deeply appreciate the time and effort they spent in attending my presentations and participating in meetings with thoughtful discussions. Their guidance and support significantly enhanced my understanding and perspective of research.

I am thankful to my family, whose support and belief in my capabilities have been a constant source of strength and motivation. A special mention goes to the newest member of our family, my unborn son, who is due to arrive in two months. Even before your arrival, you have been the joy in my life.

Jiang, Hao B.S., Economics and Mathematics, University of Kansas
M.A., Economics and Applied Statistics, UC Santa Barbara

Essays in Robust Optimization
With Applications to Finance and Renewable Energy

Advisor: Dr. Aurélie Thiele

Real-world optimization problems are often sensitive to uncertainties caused by estimation errors, forecasting inaccuracy, and imprecise information. These uncertainties bring significant challenges for decision-making in many areas. In portfolio optimization, inaccurate estimation of asset return and risk from limited historical data can cause higher risk exposure and lower expected gains for investors. For long-term energy planning problems, uncertainties in future demand, fuel prices, and other costs complicate capacity expansion. Robust optimization provides techniques to address uncertainty by planning against a "reasonable" worst-case situation.

This study of finance aims to propose a tractable robust optimization model for a Mean-Variance portfolio selection problem. We consider Markowitz's Mean-Variance Optimization when stock returns are modeled using Sharpe's single-index framework, but the model coefficients, which are denoted α and β, are not precisely known. This study assumes the α and β coefficients estimated using least-square estimators are within a prespecified ε of optimality and build a tractable robust optimization model to address this problem. The approach combines both predictive analytics in the estimation of α and β and prescriptive analytics in the optimization of the portfolio. I apply the approach to the Dow Jones and NASDAQ100 indices and find in numerical experiments that there exists an optimal $\varepsilon > 0$ for which the optimal robust portfolio achieves higher expected return and lower volatility than the benchmark portfolio obtained when $\varepsilon = 0$.

The study of renewable energy focuses on using a robust optimization approach for long-term renewable energy planning under uncertainty. The goal is to determine capacity expansion for infrastructures and electricity generation to meet state-level environmental targets, including Renewable Portfolio Standards and clean electricity requirements, over a decade-long planning horizon. The proposed model incorporates realistic factors, including the construction leading time and potential capacity for renewable energy based on geographic factors. To address the parameter uncertainty in long-term planning, we develop a tractable reformulation for the robust optimization problem. Finally, we provide numerical experiments based on real California data and provide decision-makers with various strategies for capacity investment and generation profiles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

1.1. Background

In this dissertation, we study how using robust optimization can lead the decision-maker to make better decisions in two application areas faced with high uncertainty. The first application is a portfolio selection problem based on Sharpe's single-index model. The second application is a long-term energy planning model integrating targets for renewable energy sources. In a real-world setting, uncertainty often brings challenges to the optimization of these problems.

The uncertainty can be caused by estimation errors or inaccurate predictions. An optimization problem without uncertainty is called a deterministic problem, which can be expressed as follows when it is linear:

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{1.1}$$

where Problem 1.1 is a standard form of linear programming. Parameters $c \in \mathbb{R}^{n\times 1}$, $b \in \mathbb{R}^{m\times 1}$, and $A \in \mathbb{R}^{m\times n}$ are parameters assumed known. $x \in \mathbb{R}^{n\times 1}$ is the decision variable. However, in practice, we do not have perfect knowledge of these parameters, especially in fast-changing environments or for long-term planning.

Stochastic optimization is one method proposed to address parameter uncertainty. It requires knowledge of the probability distributions for the uncertain parameters. See textbook [1] for the resource. The stochastic optimization problem can be expressed as:

$$\begin{aligned} \min \quad & \mathbb{E}_{\xi\sim\mathbb{P}}[c(\xi)^T x] \\ s.t. \quad & Pr_{\xi\sim\mathbb{P}}[A(\xi)x = b(\xi)] \geq 1-\alpha \\ & x \geq 0 \end{aligned} \tag{1.2}$$

where ξ is a random vector following a probability distribution $\mathbb{P}$. The objective is to minimize the expectation of the cost. $1-\alpha$ represents a confidence level of a guaranteed probability of the realizations of constraints. However, the probability distribution $\mathbb{P}$ is difficult to describe in a real-world setting. Challenges can also arise in the evaluation of expectation $\mathbb{E}_{\mathbb{P}}$ and probability $Pr_{\mathbb{P}}$. In addition, the formulation can be intractable.

Robust optimization is another major framework to address uncertainty. We define the uncertain parameter ξ as an element of a bounded uncertainty set $\mathcal{U}$. The concept of robust optimization is to minimize the consequences of the worst case. We refer to [2] for details. The problem can be represented as follows:

$$\begin{aligned} \min \quad \max_{\xi\in\mathcal{U}} \quad & c(\xi)^T x \\ s.t. \quad & A(\xi)x = b(\xi) \\ & x \geq 0 \end{aligned} \tag{1.3}$$

Alternatively, we can reformulate the problem as the following equivalent expression:

$$\begin{aligned} \min \quad & z \\ s.t. \quad & c^T(\xi)x \leq z \qquad \forall \xi \in \mathcal{U} \\ & A(\xi)x = b(\xi) \qquad \forall \xi \in \mathcal{U} \\ & x \geq 0 \end{aligned} \tag{1.4}$$

The optimal solution for robust optimization is against the worst-case value in the uncertainty set. The challenges include the construction of the uncertainty set, the tractability of reformulation, and over-conservatism.

The third framework is distributionally robust optimization. In this approach, ξ is defined as a random vector following a probability distribution $\mathbb{P}$, as in stochastic optimization. However, the knowledge of the probability distribution can be partially known instead of complete information. The distributionally robust optimization can be formulated as follows:

$$\begin{aligned} \min_{x} \quad \max_{\mathbb{P} \in \mathcal{P}} \quad & \mathbb{E}_{\xi \sim \mathbb{P}}[c(\xi)^T x] \\ s.t. \quad & Pr_{\xi \sim \mathbb{P}}[A(\xi)x = b(\xi)] \geq 1 - \alpha \\ & x \geq 0 \end{aligned} \tag{1.5}$$

where $\mathcal{P}$ is an uncertainty set that contains probability distributions. The concept of the approach is the combination of stochastic optimization and robust optimization that the optimal solution minimizes the expected cost against the worst-case distribution with an acceptable probability. The difficulties include the selection of the uncertainty set $\mathcal{P}$ and the tractability. We refer to [3] for the review.

The modern portfolio theory pioneered by Markowitz [4] aims to minimize the portfolio risk and ensure an acceptable level of portfolio return or maximize the return while not

exceeding a given level of risk. However, the inputs of classical Markowitz's Mean-Variance model, such as expected return and covariance matrix, must be estimated from the historical data. Therefore, estimation errors can impact the results of the portfolio optimization. In this study, we propose a robust optimization model using Sharpe's single-index model, where the model coefficients alpha and beta reflect the stock performance and sensitivity to the market, respectively. The values of alpha and beta of a stock are estimated from its historical data. In practice, the alpha and beta estimated by applying Ordinary Least Squares Estimation can partially represent the characteristics of the stock in the future due to the uncertainty. However, the actual alpha and beta driving the future random process of the stock return is close to the estimated values. Therefore, we constructed an uncertainty set that the estimated values of alpha and beta from historical data are within ε of the true values. The decision-makers choose the value of ε to reflect their viewpoints towards the magnitude of the stock can differ from the past.

Renewable energy sources have seen massive growth in recent decades. The decreasing investment costs, environmental benefits, and government policies are driving this expansion. Many states have announced their Renewable Portfolio Standards mandating minimum percentages of electricity generated from renewable energy sources. However, effectively planning the transition to higher renewable penetration faces multiple challenges. Renewable energy generation can be highly related to environmental patterns. In addition, the construction of new renewable power infrastructures requires lead times. Capacity cannot be instantly available due to years of planning, permitting, and construction preceding the operation. Furthermore, there are geographic limitations on the maximum potential renewable capacity. The uncertainty in the projections of future electricity demand, fuel prices, operations and maintenance costs, and electricity prices bring challenges to long-term energy planning. Therefore, this study develops a robust optimization model to determine state-level energy planning of capacity expansion and electricity generation profiles to meet environmental targets over multi-period horizons.

1.2. Main Contributions

The main contributions of this dissertation are the following:

- We investigate the benefits of considering a joint predictive-prescriptive optimization model in finance instead of separating the prediction and optimization phases. This is the key innovation of this model.
- We investigate the benefits of using robust optimization for long-term (decade-long) energy generation planning, which to the best of our knowledge has not been considered before. Previous works ignored important features such as the lead time needed to build a plant and bring it online and adopted a myopic approach.
- Further, we consider uncertainty on the resource cost, fixed operations and maintenance costs, variable operations and maintenance costs, electricity price and electricity demand in our energy planning model. Considering uncertainty on these parameters simultaneously for a long-term planning horizon with lead times is also novel. Finally, the portfolio of technologies available involves a very wide range of options due to numerous resource and cost classes, each representing a possible investment for the policymaker. Considering such a large number of resource and cost classes in an energy planning optimization problem is also novel. Hence, the model in this dissertation is uniquely realistic, fine-grained and complex, representing an important tool for energy policymakers.
- In each case study, we use extensive real data obtained from publicly available online sources and provide structural insights into the optimal solution. The robust solution shows more diversification than the deterministic one. These insights make our model more likely to become adopted in practice than other available techniques.

1.3. Dissertation Structure

The dissertation is organized as follows.

In Chapter 2, we propose a robust optimization approach for portfolio optimization using Sharpe's single-index model. In Section 2.1, we review the literature on portfolio optimization and different approaches addressing the uncertainty in this problem. In Section 2.2.1, we introduce the classical Markowitz's model and show how to integrate it with Sharpe's single-index model. In addition, we derive the optimal value of Sharpe's model coefficients by applying Ordinary Least Squares Estimation. In Section 2.3, we develop a robust optimization problem by considering the uncertainty in Sharpe's model coefficients and provide a tractable reformulation. In Section 2.4, we conduct numerical experiments and show the advantages of our proposed approach that provides decision-makers with trade-offs between return and risk. In Chapter 3, we develop a comprehensive long-term energy planning model integrated with goals for U.S. states regarding renewable energy generation. We apply a robust optimization framework to this problem with uncertainty. In Section 3.2, we propose a long-term energy planning model to determine capacity expansion and electricity generation profiles to meet various environmental targets. In Section 3.3, we consider the robust optimization framework of the problem and provide a tractable formulation. In Section 3.4, we perform numerical experiments using California data. In Chapter 4, we provide concluding remarks and discuss directions for future work.

Chapter 2

Finance

2.1. Introduction

Portfolio optimization under uncertainty was first proposed by Markowitz [4]. The core idea of Markowitz's Mean-Variance Optimization is to achieve at least a set expected return while minimizing the portfolio risk measured by its variance, or, to maximize the return of investment for a given risk level. In Markowitz's original model, the expected value and covariance of the stock returns are assumed known. In practice, however, theses parameters may be subject to significant uncertainty, which then can change the optimal portfolio allocation considerably [5].

Based on Markowitz's model, the Capital Asset Pricing Model (CAPM) was individually developed and refined by Treynor [6], Sharpe [7], Lintner [8] and Mossin [9]. The CAPM posits a relationship between the expected return of a stock and the market premium. In this study, the proposed model is motivated by Sharpe's single-index model where the return of each stock depends on the fluctuations of a single market factor through a simple liner regression. The coefficients of this linear regression are denoted α for the constant term and β for the coefficient about the single market factor, or slope.

Sharpe's single-index model is closely related to CAPM. The difference is that in CAPM, a stock's expected abnormal return (denoted α in Sharpe's single-index model) is zero in the long run. Black et al. [10] and Fama and MacBeth [11] provided the initial examination of CAPM, and analyzed the case when the residual returns exhibit cross-sectional correlatoin, which violates the assumption of ordinary least squares that was found in previous models.

Black et al. [10] abandoned the risk-free interest rate in CAPM and proposed a two-factor model called zero-beta CAPM. Fama and MacBeth [11] extended the CAPM into a three-factor model and showed that the model can explain more than 90% of the portfolio's returns, and introduced size and book-to-market factors called SMB (small minus big) and HML (high minus low), respectively. Subsequently, to improve the explanation of the model, Carhart [12] proposed a four-factor model, and Fama and French [13] introduced a five-factor model adding profitability and investment patterns.

Numerous theoretical frameworks have been studied to address uncertain fluctuations in the parameters of the Mean-Variance Optimization problem. Research in robust optimization (RO), pioneered by Soyster [14], and furthered in EL Ghaoui and Lebret [15], and Ben-Tal and Nemirovski [16, 17] has been widely applied in recent decades to finance problems. In particular, Goldfrab and Iyengar [18] applied RO to portfolio selection. They developed a factor model for the stock returns, and constructed robust models for Mean-Variance and Value-at-Risk problems. In addition they proved that the max-min optimization problem for the worst case could be reformulated as a second-order cone programming problem assuming that the uncertainty in the market parameters is known and bounded. Bertsimas and Sim [19] proposed a new robust optimization framework based on polyhedral uncertainty sets which led to probabilistic guarantees under mild assumptions. Extensions of robust portfolio selection have included for instance derivative guarantees [20] and copulas [21].

In typical predict-then-optimize, machine learning is used to reduce prediction error, and the prediction is used as input to the optimization problem. A Smart "Predict, then Optimize" (SPO) with a convex surrogate loss function (SPO+) for training prediction models was developed to minimize the decision error (instead of the prediction error) [22]. In their numerical experiments considering classical Markowitz's portfolio optimization with 50 synthetic assets, they compare SPO+ with Absolute Loss, Least Squares, and Random Forests methods. SPO+ outperforms Random Forests under both linear and nonlinear

models; however, the performance of SPO+ is very close to Absolute Loss and Least Squares for a linear model. The goal of a smart "Predict, then Optimize" approach is to proceed in two steps where (i) first the decision maker predicts any uncertain input parameters using a machine learning (ML) model trained on historical data, and (ii) the manager then generates decisions by solving the corresponding optimization problem using the predicted parameters.

The objective of this study is to achieve both steps simultaneously, taking into account estimation errors. A key difference between the approach in [22] for portfolio optimization and this study is that they apply the SPO+ empirical risk minimization problem to achieve a preferred risk level, whereas my approach assumes that coefficients α and β are estimated within a prespecified ε of optimality and chooses a minimum expected return, and provides the decision maker with a trade-off between the expected return and the risk level.

This study assumes that the stock returns obey Sharpe's single-index model but the coefficients α and β of the model are not known precisely. Instead, the parameters estimated using ordinary least-squares are within ε of the unknown α and β deriving the stock returns in that model, where distance is measured using the ℓ_1-distance. Then, a tractable RO model is built to find the optimal portfolio while accounting for the fact that the parameters α and β are not known precisely. ε-optimality was previously investigated by Bertsimas et al. [23], who proposed an ϵ-arbitrage approach for replicating derivative securities using stochastic dynamic programming. They found a self-financing dynamic portfolio strategy that approximates the payoff function of the option to within approximation error ϵ, where ϵ is the root-mean-squared error of an optimal-replication strategy. Their numerical experiments show that the approximation error ϵ measures market incompleteness, with ϵ equal to zero in a complete market. Later, [24] and [25] combined ϵ-arbitrage and robust optimization to solve the option pricing problem, where the uncertainty set is constructed by using historical stock prices. The advantages of their method are: tractability in pricing high-dimensional problems, and ability to model various types of options and investors' risk appetites. Nu-

merical experiments using empirical data show that prices obtained using their method are close to the observed prices in the options market. [26] provide an approximate optimality theorem and approximate duality theorems on approximate solutions (quasi ϵ-solutions) for a robust convex programming problem under data uncertainty.

Contributions:

The goal of this study is to investigate a joint predictive-prescriptive analytics approach in Sharpe's single-index model, both from a standpoint of tractability of the robust reformulation and performance in numerical experiments. Instead of first estimating key parameters using ordinary least squares and then adding a robust modelization of the uncertain parameters centered at their estimated values to optimize the portfolio, this study investigates the benefits of combining both steps in a single formulation. This is the key novel feature of the work.

In Section 2.2 we present Sharpe's single-index model without uncertainty. In Section 2.3 we introduce uncertainty on the parameters of the single-index model, where we assume that their estimated values are within ε of the true values, which are unknown, and we develop a tractable reformulation of this predictive-prescriptive analytics model. In Section 2.4 we present numerical results. Those experiments empirically support the existence of good choices for ε. Further, in those experiments, our approach leads to portfolio allocations that, for the expected return and volatility of simulated data, outperform robust portfolio allocations from the literature [18] and thus augment the decision-maker's menu of portfolios he can choose from based on his attitudes toward risk and return. In section 2.5, we summarize the chapter and discuss the future study.

2.2. Sharpe's Single-Index Model without Parameter Uncertainty

In this section, a review of classical Markowitz's Mean-Variance Optimization is presented when stock returns follow Sharpe's single-index model. The optimal Sharpe's α and β coefficients are derived by applying Ordinary Least Squares Estimation (OLS). The α coefficient is the constant and β is the slope in the simple linear regression explaining the stock returns as a function of the single index.

2.2.1. Classical Markowitz's Model

The following notation is used.

N: the number of stocks in the portfolio,

μ_i: the average return of stock i for all i,

q_{ij}: the covariance between stock i and stock j for all i, j,

w: the minimum expected return from the portfolio,

x_i: the fraction of the portfolio invested in stock i.

The traditional Markowitz model, in which the decision maker aims to minimize portfolio variance as a measure of risk while ensuring that the expected return meets or exceeds a

certain threshold (with no short selling allowed), is defined as follows:

$$\min_{x} \sum_{i=1}^{N}\sum_{j=1}^{N} q_{ij}x_i x_j \qquad \textbf{(Problem 1)}$$

$$\text{s.t.} \sum_{i=1}^{N} \mu_i x_i \geq w \tag{2.1}$$

$$\sum_{i=1}^{N} x_i = 1 \tag{2.2}$$

$$x_i \geq 0 \quad \forall\, i \tag{2.3}$$

In the absence of uncertainty, the expected return vector μ and covariance matrix Q are estimated by fitting a model (such as a time series model) to historical data. However, these estimates are subject to uncertainty due to the possibility that future stock price behavior may not be accurately represented by past data. Before modeling this uncertainty, we describe the Sharpe's single-factor model below.

2.2.2. The Single-Index Model

This section investigates Markowitz's problem under the assumption that stock returns follow Sharpe's single-index model. According to this model, the return of stock i in time period t is modeled as:

$$r_{it} = \alpha_i + \beta_i r_{mt} + e_{it} \tag{2.4}$$

The following notation is used:

r_{it} : return of stock i in time period t,

r_{mt} : return of market index in time period t,

α_i : expected return of stock i, which is independent of market performance,

β_i : Sharpe's beta, which measures how sensitive a stock is to market moves,

e_{it} : the residual return of r_{it}, which is assumed to be $i.i.d.$ with mean zero,

σ_{ei}^2 : variance of the residual return e_{it} for stock i,

σ_m^2 : variance of market index.

The average return of stock i from time period 1 to T can be written as:

$$\hat{\mu}_i = \frac{1}{T}\sum_{t=1}^{T} r_{it}. \tag{2.5}$$

The average return on the market index from time period 1 to T can be written as:

$$\hat{\mu}_m = \frac{1}{T}\sum_{t=1}^{T} r_{mt}. \tag{2.6}$$

Therefore, Eq. (2.4) can be reformulated as:

$$\hat{\mu}_i = \alpha_i + \beta_i \hat{\mu}_m. \tag{2.7}$$

The variance of stock i is given by:

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{ei}^2. \tag{2.8}$$

Then, the covariance between stock i and stock j in timer period it is:

$$\begin{aligned} Cov(r_{it}, r_{jt}) &= E[(r_{it} - \hat{\mu}_i)(r_{jt} - \hat{\mu}_j)] \\ &= E[(\beta_i(r_{mt} - \hat{\mu}_m) + e_{it})(\beta_j(r_{mt} - \hat{\mu}_m) + e_{jt})] \\ &= \beta_i\beta_j E[(r_{mt} - \hat{\mu}_m)^2] \\ &= \beta_i\beta_j\sigma_m^2 \end{aligned} \tag{2.9}$$

Parameters α_i and β_i for all i in the simple linear regression can be computed by using Ordinary Least Squares Estimation (OLS). In OLS, the decision-maker seeks to find the optimal values of α_i and β_i that minimize the residual sum of squares

$$S(\alpha_i, \beta_i) = \frac{1}{T}\sum_{t=1}^{T}(\alpha_i + \beta_i r_{mt} - r_{it})^2. \tag{2.10}$$

Lemma 2.1 *Given the parameters $T, \hat{\mu}_m, \hat{\mu}_i, r_{mt}$, and r_{it}, the optimal α_i^* and β_i^* minimizing the residual sum of squares satisfy the following:*

$$\bar{\alpha}_i^* = \frac{\hat{\mu}_i(\sum_{t=1}^{T} r_{mt}^2) - \hat{\mu}_m(\sum_{t=1}^{T} r_{mt}r_{it})}{(\sum_{t=1}^{T} r_{mt}^2) - T\hat{\mu}_m^2}, \tag{2.11}$$

$$\bar{\beta}_i^* = \frac{(\sum_{t=1}^{T} r_{mt}r_{it}) - T\hat{\mu}_i^2}{(\sum_{t=1}^{T} r_{mt}^2) - T\hat{\mu}_m^2}. \tag{2.12}$$

Proof: This follows from setting the gradient of this quadratic, convex function to zero.

We solve for all i:

$$\frac{\partial S}{\partial \alpha_i} = 0 \;\Rightarrow\; \frac{2}{T}\sum_{t=1}^{T}(\alpha_i + \beta_i r_{mt} - r_{it}) = 0$$
$$\frac{\partial S}{\partial \beta_i} = 0 \;\Rightarrow\; \frac{2}{T}\sum_{t=1}^{T} r_{mt}(\alpha_i + \beta_i r_{mt} - r_{it}) = 0$$

This can be rewritten as, for all i:

$$\begin{cases} \alpha_i + \beta_i \hat{\mu}_m - \hat{\mu}_i = 0 \\ \alpha_i \hat{\mu}_m + \beta_i \dfrac{1}{T}\displaystyle\sum_{t=1}^{T} r_{mt}^2 - \dfrac{1}{T}\sum_{t=1}^{T} r_{mt} r_{it} = 0 \end{cases}$$

Solving this system of equations yields the model without uncertainty for all i, resulting $\bar{\alpha}_i^*$ and $\bar{\beta}_i^*$.

□

2.3. Proposed Portfolio Selection Model

In practice, there is uncertainty on the model parameters, and we are interested in understanding how the optimal portfolio allocation will be impacted if the estimated values of α and β using historical data are not the true values but are within ε of the true values, as measured by the ℓ_1-Norm. This model combines both predictive and prescriptive analytics by jointly estimating key model parameters and optimizing the portfolio. The uncertainty set is motivated by the idea that the actual α and β driving the future random process of stock returns are likely to be close (but not equal) to the estimated values calculated using historical data, potentially due to unmodeled factors. No assumption is made on the distribution of the error in the model, and ε is a hyperparameter selected by the decision-maker

to capture their opinion of how much the future random process may differ from the past. The concept of ε-optimality is formalized in Definition 2.1 below.

Definition 2.1 $\varepsilon - optimality$. We say that a solution is within ε of optimality if the ℓ_1-Norm of the gradient of the objective in the ordinary least squares problem is at most ε. □

Definition 2.1 implies that α and β satisfy, at ε of optimality:

$$\sum_{i=1}^{N}[|\frac{1}{T}\sum_{t=1}^{T}(\alpha_i+\beta_i r_{mt}-r_{it})|+|\frac{1}{T}\sum_{t=1}^{T}r_{mt}(\alpha_i+\beta_i r_{mt}-r_{it})|]\leq\varepsilon. \tag{2.13}$$

To linearize Eq. (2.13), let $y_i=|\frac{1}{T}\sum_{t=1}^{T}(\alpha_i+\beta_i r_{mt}-r_{it})|$ and $z_i=|\frac{1}{T}\sum_{t=1}^{T}r_{mt}(\alpha_i+\beta_i r_{mt}-r_{it})|$, and then can be reformulated as:

$$-y_i\leq\frac{1}{T}\sum_{t=1}^{T}(\alpha_i+\beta_i r_{mt}-r_{it})\leq y_i,\qquad\forall i \tag{2.14}$$

$$-z_i\leq\frac{1}{T}\sum_{t=1}^{T}r_{mt}(\alpha_i+\beta_i r_{mt}-r_{it})\leq z_i\qquad\forall i \tag{2.15}$$

Let $\mathcal{S}_\varepsilon=\Bigg\{(\alpha,\beta)\Bigg|\exists\ y,z\geq 0,\ \sum_{i=1}^{N}(y_i+z_i)\leq\varepsilon,\ -y_i\leq\frac{1}{T}\sum_{t=1}^{T}(\alpha_i+\beta_i r_{mt}-r_{it})\leq y_i,$
$-z_i\leq\frac{1}{T}\sum_{t=1}^{T}[r_{mt}(\alpha_i+\beta_i r_{mt}-r_{it})]\leq z_i\ \forall i\Bigg\}.$

Then, **Problem 1** with parameter uncertainty can be reformulated as:

$$\min_{x} \quad \max_{\alpha,\beta\in\mathcal{S}_\varepsilon} \sigma_m^2(\sum_{i=1}^{N} x_i\beta_i)^2 \tag{**Problem 2**}$$

$$\text{s.t.} \quad \min_{\tilde{\alpha},\tilde{\beta}\in\mathcal{S}_\varepsilon} \sum_{i=1}^{N}(\tilde{\alpha}_i + \tilde{\beta}_i\hat{\mu}_m)x_i \geq w \tag{2.16}$$

$$\sum_{i=1}^{N} x_i = 1 \tag{2.17}$$

$$x_i \geq 0 \quad \forall i \tag{2.18}$$

The inner problem of maximizing a convex function in the objective of **Problem 2** can be rewritten as two tractable sub-problems. Introduce a new variable $v = |\sum_i x_i\beta_i|$, so that the objective function can be formulated as:

$$\min_{x} \quad \max_{\alpha,\beta\in S_\varepsilon} \sigma_m^2 v^2$$

Theorem 2.2 *Sharpe's single-index model with parameter uncertainty **Problem 2** can be reformulated as the following tractable Robust Problem (**RP**):*

$$\begin{aligned}
\min \quad & \sigma_m^2 v^2 && \text{(**RP**)}\\
s.t. \quad & \varepsilon p_1 + \sum_i \hat{\mu}_i(p_{2i} - p_{3i}) + \frac{1}{T}\sum_i\sum_t r_{mt}r_{it}(p_{4i} - p_{5i}) \leq v\\
& p_{2i} - p_{3i} + \hat{\mu}_m(p_{4i} - p_{5i}) = 0 \quad \forall i\\
& \hat{\mu}_m(p_{2i} - p_{3i}) + \frac{1}{T}\sum_t r_{mt}^2(p_{4i} - p_{5i}) = x_i \quad \forall i\\
& p_1 - p_{2i} - p_{3i} \geq 0 \quad \forall i\\
& p_1 - p_{4i} - p_{5i} \geq 0 \quad \forall i\\
& p_1,\ p_{2i},\ p_{3i},\ p_{4i},\ p_{5i} \geq 0\ \forall i
\end{aligned}$$

$$\varepsilon k_1 + \sum_i \hat{\mu}_i(k_{2i} - k_{3i}) + \frac{1}{T}\sum_i \sum_t r_{mt} r_{it}(k_{4i} - k_{5i}) \leq v$$

$$k_{2i} - k_{3i} + \hat{\mu}_m(k_{4i} - k_{5i}) = 0 \quad \forall i$$

$$\hat{\mu}_m(k_{2i} - k_{3i}) + \frac{1}{T}\sum_t r_{mt}^2(k_{4i} - k_{5i}) = -x_i \quad \forall i$$

$$k_1 - k_{2i} - k_{3i} \geq 0 \quad \forall i$$

$$k_1 - k_{4i} - k_{5i} \geq 0 \quad \forall i$$

$$k_1,\ k_{2i},\ k_{3i},\ k_{4i},\ k_{5i} \geq 0\ \forall i$$

$$\varepsilon q_1 + \sum_i \hat{\mu}_i(q_{2i} - q_{3i}) + \frac{1}{T}\sum_i \sum_t r_{mt} r_{it}(q_{4i} - q_{5i}) \geq w$$

$$q_{2i} - q_{3i} + \hat{\mu}_m(q_{4i} - q_{5i}) = x_i \quad \forall i$$

$$\hat{\mu}_m(q_{2i} - q_{3i}) + \frac{1}{T}\sum_t r_{mt}^2(q_{4i} - q_{5i}) = \hat{\mu}_m x_i \quad \forall i$$

$$q_1 - q_{2i} - q_{3i} \leq 0 \quad \forall i$$

$$q_1 - q_{4i} - q_{5i} \leq 0 \quad \forall i$$

$$q_1,\ q_{2i},\ q_{3i},\ q_{4i},\ q_{5i} \leq 0\ \forall i$$

$$\sum_i x_i = 1$$

$$x_i \geq 0 \quad \forall i$$

Note that σ_m^2 in the objective function can be omitted. Then the objective is equivalent to min v^2.

Proof: This is a direct application of strong duality in linear programming. **Problem 2** can be rewritten as:

$$\min_{x,v}\ \sigma_m^2 v^2 \qquad \textbf{(Problem 3)}$$

$$\text{s.t. } v \geq \max_{\alpha,\beta \in \mathcal{S}_\varepsilon} \sum_{i=1}^{N} x_i \beta_i$$

$$
\begin{aligned}
& v \geq \max_{\alpha,\beta \in \mathcal{S}_\varepsilon} -\sum_{i=1}^{N} x_i \beta_i \\
& \min_{\tilde{\alpha},\tilde{\beta} \in \mathcal{S}_\varepsilon} \sum_{i=1}^{N} (\tilde{\alpha}_i + \tilde{\beta}_i \hat{\mu}_m) x_i \geq w \\
& \sum_{i=1}^{N} x_i = 1 \\
& x_i \geq 0 \quad \forall i.
\end{aligned}
$$

The three inner problems

$$\max_{\alpha,\beta \in \mathcal{S}_\varepsilon} \sum_{i=1}^{N} x_i \beta_i \tag{P1}$$

$$\max_{\alpha,\beta \in \mathcal{S}_\varepsilon} -\sum_{i=1}^{N} x_i \beta_i \tag{P2}$$

$$\min_{\tilde{\alpha},\tilde{\beta} \in \mathcal{S}_\varepsilon} \sum_{i=1}^{N} (\tilde{\alpha}_i + \tilde{\beta}_i \hat{\mu}_m) x_i \tag{P3}$$

are linear problems over non-empty, bounded feasible sets; thus, **Problem 3** can be reformulated in a tractable manner using strong duality in linear programming, as in [19]. To formulate the dual problem, rewrite $\mathcal{S}_\varepsilon$ as:

$$
\begin{aligned}
& \sum_i (y_i + z_i) \leq \varepsilon \\
& \alpha_i + \hat{\mu}_m \beta_i - y_i \leq \hat{\mu}_i \quad \forall i \\
& -\alpha_i - \hat{\mu}_m \beta_i - y_i \leq -\hat{\mu}_i \quad \forall i \\
& \hat{\mu}_m \alpha_i + (\frac{1}{T} \sum_t r_{mt}^2) \beta_i - z_i \leq \frac{1}{T} \sum_t r_{mt} r_{it} \quad \forall i \\
& -\hat{\mu}_m \alpha_i - (\frac{1}{T} \sum_t r_{mt}^2) \beta_i - z_i \leq -\frac{1}{T} \sum_t r_{mt} r_{it} \quad \forall i \\
& y_i \geq 0, z_i \geq 0 \quad \forall i.
\end{aligned}
$$

The dual problem of the linear problem (P1), (P1), and (P3) are, respectively:

$$\min_{p} \quad \varepsilon p_1 + \sum_i \hat{\mu}_i(p_{2i} - p_{3i}) + \frac{1}{T}\sum_i \sum_t r_{mt} r_{it}(p_{4i} - p_{5i}) \qquad \textbf{(Dual 1)}$$

$$\begin{aligned}
\text{s.t.} \quad & p_{2i} - p_{3i} + \hat{\mu}_m(p_{4i} - p_{5i}) = 0 \quad \forall i \\
& \hat{\mu}_m(p_{2i} - p_{3i}) + \frac{1}{T}\sum_t r_{mt}^2(p_{4i} - p_{5i}) = x_i \quad \forall i \\
& p_1 - p_{2i} - p_{3i} \geq 0 \quad \forall i \\
& p_1 - p_{4i} - p_{5i} \geq 0 \quad \forall i \\
& p_1,\ p_{2i},\ p_{3i},\ p_{4i},\ p_{5i} \geq 0 \quad \forall i.
\end{aligned}$$

$$\min_{k} \quad \varepsilon k_1 + \sum_i \hat{\mu}_i(k_{2i} - k_{3i}) + \frac{1}{T}\sum_i \sum_t r_{mt} r_{it}(k_{4i} - k_{5i}) \qquad \textbf{(Dual 2)}$$

$$\begin{aligned}
\text{s.t.} \quad & k_{2i} - k_{3i} + \hat{\mu}_m(k_{4i} - k_{5i}) = 0 \quad \forall i \\
& \hat{\mu}_m(k_{2i} - k_{3i}) + \frac{1}{T}\sum_t r_{mt}^2(k_{4i} - k_{5i}) = -x_i \quad \forall i \\
& k_1 - k_{2i} - k_{3i} \geq 0 \quad \forall i \\
& k_1 - k_{4i} - k_{5i} \geq 0 \quad \forall i \\
& k_1,\ k_{2i},\ k_{3i},\ k_{4i},\ k_{5i} \geq 0 \ \forall i
\end{aligned}$$

$$\max_{q} \quad \varepsilon q_1 + \sum_i \hat{\mu}_i(q_{2i} - q_{3i}) + \frac{1}{T}\sum_i \sum_t r_{mt} r_{it}(q_{4i} - q_{5i}) \qquad \textbf{(Dual 3)}$$

$$\begin{aligned}
\text{s.t.} \quad & q_{2i} - q_{3i} + \hat{\mu}_m(q_{4i} - q_{5i}) = x_i \quad \forall i \\
& \hat{\mu}_m(q_{2i} - q_{3i}) + \frac{1}{T}\sum_t r_{mt}^2(q_{4i} - q_{5i}) = \hat{\mu}_m x_i \quad \forall i \\
& q_1 - q_{2i} - q_{3i} \leq 0 \quad \forall i \\
& q_1 - q_{4i} - q_{5i} \leq 0 \quad \forall i \\
& q_1,\ q_{2i},\ q_{3i},\ q_{4i},\ q_{5i} \leq 0 \quad \forall i
\end{aligned}$$

The dual problems are derived by reinjecting the dual formulations into **Problem 3** and using, as in [19], the fact that only need to find feasible dual variables for the constraints to be satisfied (the minimizations in Dual 1 and Dual 2 and the maximization in Dual 3 can be dropped), resulting in the tractable reformulation presented in Theorem 2.2. □

2.4. Numerical Experiments

In this section, the results of numerical experiments after solving the proposed **RP** are analyzed. The purpose of these experiments is to examine the potential benefits of addressing the Sharpe's single-index model under parameter uncertainty using a joint predictive-prescriptive analytics approach. These benefits may include reduced portfolio risk, increased portfolio return, and greater insights into portfolio allocation. Specifically, the risk and return profile of the robust optimal allocation is compared to the benchmark ($\varepsilon = 0$). Two types of numerical experiments are conducted.

The first experiment compares the structure of portfolios generated by solving the **RP** using real-world market indices for different values of w and ε, with w and ε varied to see how the portfolio structure changes. The second numerical experiment investigates the performance of optimal robust portfolio on simulated data when the true α and β are random and computed by adding a random term to the values estimated using real-world market indices. The portfolio return, calculated as $\sum_{i=1}^{N} r_i x_i$, and the volatility of simulated data under different values of ε are compared to the benchmark ($\varepsilon = 0$) for given levels of w. All the numerical experiments are conducted using the NEOS Server for CPLEX/AMPL and R version 4.0.1 on the platform:

Model	Apple Macbook Pro 2018
Processor	Intel 6-core CPU 2.2 GHz
RAM	32 GB

2.4.1. Portfolios Structures for Real Market Indices

This section compares the portfolios structures of our model given different scenarios of w and ε to the benchmark ($\varepsilon = 0$) using real market indices, namely the Dow Jones Industrial Average (DJIA) and NASDAQ 100 Index (NDX). Dow Jones quantifies the stock performance of thirty large representative companies in the United States, and NASDAQ 100 consists of one hundred of the largest non-financial companies.

The ten-year monthly DJIA and NDX data from March 1, 2009 to February 1, 2019 is considered. To maintain the continuity of the price-weighted DJIA, an index divisor is applied whenever the composition of the index changes. During this period, several companies were added, replaced, spun off, or merged into the DJIA and NDX. The latest updated thirty components of the DJIA as of April 2020 were selected for the data set, but the Dow Chemical Company (DOW Inc.) was excluded because it was not included in the index between 2009 and 2019. As a result, the data set using the DJIA includes 29 stocks and 119 time periods. The same procedure was used to select stocks from the NDX from March 1, 2009 to February 1, 2019, resulting in a data set with 86 components ($n = 86$) available for the NDX in this time period (see Table 2.1).

A fixed Dow divisor was used on February 1, 2019 to compute the modified DJIA corresponding to the twenty-nine components by dividing the DJIA adjusted closing price by the summation of the twenty-nine components' adjusted closing prices. The original NASDAQ 100 is a capitalization-weighted index, and the same method was used to calculate a price-weighted index corresponding to the selected 86 components from the NASDAQ 100.

2.4.1.1. Procedures to Find Portfolios Structures

To solve the robust model **RP** given different values of w and ε, the following steps are applied:

Table 2.1: Components of selected DJI 29 and NDX 86

DJI 29

AAPL	Apple Inc.	INTC	Intel Corp.	PG	The Procter & Gamble Company
AXP	American Express Company	JNJ	Johnson & Johnson	RTX	Raytheon Technologies
BA	The Boeing Company.	JPM	JPMorgan Chase & Co.	TRV	The Travelers Companies, Inc.
CAT	Caterpillar Inc.	KO	The Coca-Cola Company	UNH	UnitedHealth Group Inc.
CSCO	Cisco Systems, Inc.	MCD	McDonald's Corp.	V	Visa Inc.
CVX	Chevron Corp.	MMM	3M Commay	VZ	Verizon Communications, Inc.
DIS	The Walt Disney Company	MRK	Merk & Co., Inc.	WBA	Walgreens Boots Alliance, Inc.
GS	The Goldman Sachs Group, Inc.	MSFT	Microsoft Corp.	WMT	Walmart Inc.
HD	The Home Depot, Inc.	NKE	Nike, Inc.	XOM	Exxon Mobil Corp.
IBM	IBM Corp.	PFE	Pfizer Inc.		

NDX 86

AAPL	Apple Inc.	CTXS	Citrix Systems, Inc.	MXIM	Maxim Integrated Products, Inc.
ADBE	Adobe Inc.	DLTR	Dollar Tree Inc.	NFLX	Netflix, Inc.
ADI	Analog Devices, Inc.	DXCM	DexCom, Inc.	NTAP	NetApp, Inc.
ADP	Automatic Data Processing, Inc.	EA	Electronic Arts Inc.	NTES	NetEase, Inc.
ADSK	Autodesk, Inc.	EBAY	eBay Inc.	NVDA	NVIDIA Corp.
ALGN	Align Technology, Inc.	EXC	Exelon Corp.	ORLY	O'Reilly Automotive, Inc.
ALXN	Alexion Pharmaceuticals, Inc.	EXPE	Expedia Group, Inc.	PAYX	Paychex, Inc.
AMAT	Applied Materials, Inc.	FAST	Fastenal Company	PCAR	PACCAR Inc.
AMD	Advanced Micro Devices, Inc.	FISV	Fiserv, Inc.	PEP	PepsiCo, Inc.
AMGN	Amgen Inc.	GILD	Gilead Sciences, Inc.	QCOM	QUALCOMM
AMZN	Amazon.com, Inc.	GOOGL	Alphabet Inc.	REGN	Regeneron Pharmaceuticals, Inc.
ANSS	ANSYS, Inc.	IDXX	IDEXX Laboratories, Inc.	ROST	Ross Stores, Inc.
ASML	ASML Holding N.V.	ILMN	Illumina, Inc.	SBUX	Starbucks Corp.
ATVI	Activision Blizzard, Inc.	INCY	Incyte Corp.	SGEN	Seattle Genetics, Inc.
BIDU	Baidu, Inc. ADS	INTC	Intel Corp.	SIRI	Sirius XM Holdings Inc.
BIIB	Biogen Inc.	INTU	Intuit Inc.	SNPS	Synopsys, Inc.
BKNG	Booking Holdings Inc.	ISRG	Intuitive Surgical, Inc.	SWKS	Skyworks Solutions, Inc.
BMRN	BioMarin Pharmaceutical Inc.	KLAC	KLA Corp.	TCOM	Trip.com Group Limited
CDNS	Cadence Design Systems, Inc.	LBTYA	Liberty Global plc Class A	TMUS	T-Mobile US, Inc.
CERN	Cerner Corp.	LBTYK	Liberty Global plc Class C	TTWO	Take-Two Interactive Software, Inc.
CHKP	Check Point Software Technologies	LRCX	Lam Research Corp.	TXN	Texas Instruments
CMCSA	Comcast Corp. Class A	LULU	lululemon athletica inc.	ULTA	Ulta Beauty, Inc.
COST	Costco Wholesale Corp.	MAR	Marriott International	VRSN	VeriSign, Inc.
CPRT	Copart, Inc. (DE)	MCHP	Microchip Technology	VRTX	Vertex Pharmaceuticals
CSCO	Cisco Systems, Inc.	MDLZ	Mondelez International, Inc.	WBA	Walgreens Boots Alliance, Inc.
CSGP	CoStar Group, Inc.	MELI	MercadoLibre, Inc.	WDC	Western Digital Corp.
CSX	CSX Corp.	MNST	Monster Beverage Corp.	XEL	Xcel Energy Inc.
CTAS	Cintas Corp.	MSFT	Microsoft Corp.	XLNX	Xilinx, Inc.
CTSH	Cognizant Technology Solutions	MU	Micron Technology, Inc.		

i. Compute the returns on stocks r_{it}, returns on the market index r_{mt}, and variance of the market index σ^2.

ii. Using Eqs. (2.5) and (2.6) to calculate the average return on stock $\hat{\mu}_i$ and the average return on the market index $\hat{\mu}_m$, respectively.

iii. Choose values of the expected return w from the portfolio, and values of ε:

   a. Consider scenarios of $w = 0$, $w = 0.01$, $w = 0.015$ and $w = 0.02$.

   b. Under each scenario of w, vary ε from 0 to 0.1 with stepsize 0.0001.

iv. Solve the robust model **RP** for the combinations of w and ε.

2.4.1.2. Results for Portfolios Structures

Tables 2.2 and 2.4 document the portfolios structures (i.e., the x_i values for all i) for DJI 29, given $w = 0$ and $w = 0.01$, and $w = 0.02$, respectively. Since **RP** generates the exact same portfolio structure for $w = 0$ and $w = 0.01$, those outputs are summarized in Table 2.2. Because of page length limitations, not all ε are included, and results are presented where the number of stocks invested in changes. For instance, for DJI 29 ($w = 0$ and 0.01), varying ε from 0.0001 to 0.0003 (respectively, from 0.0004 to 0.0008) leads to the same portfolio allocation consisting of four (respectively, five) stocks, so the results for $\varepsilon = 0.001$ and 0.0008 in Table 2.2 are shown. This is why the values of ε shown change for DJI 29 and NDX 86.

Table 2.5: Portfolios Allocations for NDX 86 ($w = 0$)

i \ ε	**0 (Benchmark)**	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
AAPL	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ADBE	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ADI	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
ADP	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
ALXN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
AMAT	0	0	0	0	0	0	0	0	0	0	0.0128
AMGN	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
AMZN	0	0	0	0	0	0	0	0	0	0.0149	0.0128
ANSS	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ASML	0	0	0	0	0	0	0	0	0	0.0149	0.0128
ATVI	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
BIIB	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
BMRN	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
CDNS	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CERN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CHKP	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
CMCSA	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
COST	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
CPRT	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
CSCO	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CSGP	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
CSX	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CTAS	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
CTSH	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CTXS	0	0	0	0	0	0	0	0	0	0	0.0128
DLTR	0	0	0	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
DXCM	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
EA	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
EBAY	0	0	0	0	0	0	0	0	0	0.0149	0.0128
EXC	1	0.5	0.3333	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
EXPE	0	0	0	0	0	0	0	0	0	0	0.0128
FAST	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
FISV	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
GILD	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
GOOGL	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
IDXX	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
ILMN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
INTC	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
INTU	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
ISRG	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
KLAC	0	0	0	0	0	0	0	0	0	0	0.0128
LBTYA	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
LBTYK	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
LRCX	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
LULU	0	0	0	0	0	0	0	0	0	0.0149	0.0128
MAR	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
MCHP	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
MDLZ	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
MNST	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
MSFT	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
MU	0	0	0	0	0	0	0	0	0	0	0.0128
MXIM	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
NFLX	0	0	0	0	0	0	0	0	0	0	0.0128
NTAP	0	0	0	0	0	0	0	0	0	0	0.0128
NTES	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
NVDA	0	0	0	0	0	0	0	0	0	0	0.0128
ORLY	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
PAYX	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
PCAR	0	0	0	0	0	0	0	0	0	0.0149	0.0128
PEP	0	0	0.3333	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
QCOM	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
REGN	0	0	0	0	0	0	0	0	0	0	0.0128
ROST	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
SBUX	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
SIRI	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
SNPS	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
SWKS	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
TCOM	0	0	0	0	0	0	0	0	0	0	0.0128
TMUS	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
TTWO	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
TXN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ULTA	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
VRSN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
VRTX	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
WBA	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
WDC	0	0	0	0	0	0	0	0	0	0	0.0128
XEL	0	0.5	0.3333	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
XLNX	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128

Table 2.2: Portfolios Allocations for DJI 29 ($w = 0$ and $w = 0.01$)

i \ ε	**0 (Benchmark)**	0.0001	0.0008	0.0009	0.002	0.007	0.011	0.023	0.029	...	0.1
AAPL	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
AXP	0	0	0	0	0	0	0	0	0.0345	0.0345	0.0345
BA	0	0	0	0	0	0	0	0.0385	0.0345	0.0345	0.0345
CAT	0	0	0	0	0	0	0	0	0.0345	0.0345	0.0345
CSCO	0	0	0	0	0	0	0	0.0385	0.0345	0.0345	0.0345
CVX	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
DIS	0	0	0	0	0	0	0.0455	0.0385	0.0345	0.0345	0.0345
GS	0	0	0	0	0	0	0	0	0.0345	0.0345	0.0345
HD	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
IBM	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
INTC	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
JNJ	0	0	0	0	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
JPM	0	0	0	0	0	0	0	0.0385	0.0345	0.0345	0.0345
KO	0	0	0	0.1667	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
MCD	1	0.25	0.2	0.1667	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
MMM	0	0	0	0	0	0	0.0455	0.0385	0.0345	0.0345	0.0345
MRK	0	0	0.2	0.1667	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
MSFT	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
NKE	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
PFE	0	0	0	0	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
PG	0	0.25	0.2	0.1667	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
RTX	0	0	0	0	0	0	0.0455	0.0385	0.0345	0.0345	0.0345
TRV	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
UNH	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
V	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
VZ	0	0.25	0.2	0.1667	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
WBA	0	0	0	0	0	0	0	0.0385	0.0345	0.0345	0.0345
WMT	0	0.25	0.2	0.1667	0.125	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345
XOM	0	0	0	0	0	0.0526	0.0455	0.0385	0.0345	0.0345	0.0345

Table 2.3: Portfolios Allocations for DJI 29 ($w = 0.015$)

i \ ε	**0 (Benchmark)**	0.0001	0.0005	0.0006	0.0008	0.001	0.005	0.014	0.03	0.07	0.08 (infeasible)
AAPL	0	0	0	0	0	0	0.0678	0.0422	0.0379	0.0655	0.0834
AXP	0	0	0	0	0	0	0	0.0422	0.0379	0.0655	0.0829
BA	0	0	0	0	0	0	0	0.0422	0.0379	0.0655	0.0845
CAT	0	0	0	0	0	0	0	0	0.0379	0.0655	0.0846
CSCO	0	0	0	0	0	0	0	0	0.0379	0.0174	0.0047
CVX	0	0	0	0	0	0	0	0.0422	0.0379	0	0.0006
DIS	0	0	0	0	0	0	0	0.0422	0.0379	0.0655	0.0805
GS	0	0	0	0	0	0	0	0	0	0	0.0006
HD	0	0	0	0	0	0	0.0678	0.0422	0.0379	0.0655	0.0833
IBM	0	0	0	0	0	0	0	0.0299	0.0137	0	0.0005
INTC	0	0	0	0	0	0	0.0678	0.0422	0.0379	0.0655	0.0572
JNJ	0	0	0	0	0	0	0.0678	0.0422	0.0379	0	0.0014
JPM	0	0	0	0	0	0	0	0	0.0379	0.0655	0.0794
KO	0	0	0	0	0	0	0.0678	0.0422	0.0379	0	0.0010
MCD	0.8568	0.3813	0.2320	0.1833	0.1496	0.1368	0.0678	0.0422	0.0379	0	0.0032
MMM	0	0	0	0	0	0	0	0.0422	0.0379	0.0655	0.0797
MRK	0	0	0.2320	0.1833	0.1496	0.1368	0.0678	0.0422	0.0379	0.0655	0.0040
MSFT	0	0	0	0	0	0	0	0.0422	0.0379	0.0655	0.0848
NKE	0	0	0	0	0	0.0424	0.0678	0.0422	0.0379	0.0655	0.0830
PFE	0	0	0	0	0	0	0.0678	0.0422	0.0379	0.0655	0.0043
PG	0	0	0.0720	0.1833	0.1496	0.1368	0.0678	0.0422	0.0379	0	0.0007
RTX	0	0	0	0	0	0	0	0.0422	0.0379	0	0.0020
TRV	0	0	0	0	0	0	0	0.0422	0.0379	0	0.0029
UNH	0.1432	0.2373	0.2320	0.1833	0.1496	0.1368	0.0678	0.0422	0.0379	0.0655	0.0829
V	0	0	0	0.0833	0.1496	0.1368	0.0678	0.0422	0.0379	0.0655	0.0829
VZ	0	0.3813	0.2320	0.1833	0.1496	0.1368	0.0678	0.0422	0.0379	0	0.0014
WBA	0	0	0	0	0	0	0	0	0.0379	0	0.0025
WMT	0	0	0	0	0.1027	0.1368	0.0678	0.0422	0.0379	0	0.0005
XOM	0	0	0	0	0	0	0.0508	0.0422	0	0	0.0009

Table 2.4: Portfolios Allocations for DJI 29 ($w = 0.02$)

i \ ε	**0 (Benchmark)**	0.0001	0.0004	0.0005	0.0009	0.001	0.008	0.014	0.02	0.023	0.024 (infeasible)
AAPL	0	0	0	0	0.1678	0.1682	0.0905	0.0907	0.0988	0.1235	0.2009
AXP	0	0	0	0	0	0	0.0905	0.0907	0.0988	0.1235	0.1874
BA	0	0	0	0	0	0	0.0905	0.0907	0.0988	0.1235	0.2088
CAT	0	0	0	0	0	0	0	0.0907	0.0988	0.1235	0.1076
CSCO	0	0	0	0	0	0	0	0	0	0	0.0042
CVX	0	0	0	0	0	0	0	0	0	0	0.0098
DIS	0	0	0	0	0	0	0.0905	0.0907	0.0988	0	0.0053
GS	0	0	0	0	0	0	0	0	0	0	0.0099
HD	0	0	0	0	0	0	0.0905	0.0907	0.0988	0.1235	0.1591
IBM	0	0	0	0	0	0	0	0	0	0	0.0012
INTC	0	0	0	0	0	0	0	0	0.0120	0	0.0282
JNJ	0	0	0	0	0	0	0	0	0	0	0.0115
JPM	0	0	0	0	0	0	0	0	0	0	0.0003
KO	0	0	0	0	0	0	0	0	0	0	0.0117
MCD	0.3489	0.3422	0.3276	0.2329	0.1678	0.1682	0.0905	0.0907	0	0	0.0045
MMM	0	0	0	0	0	0	0	0	0	0	0.0276
MRK	0	0	0	0.0682	0.1611	0.1588	0.0905	0.0026	0	0	0.0038
MSFT	0	0	0	0	0	0	0.0905	0.0907	0.0988	0.0121	0.0801
NKE	0	0	0.0173	0.2329	0.1678	0.1682	0.0905	0.0907	0.0988	0.1235	0.1746
PFE	0	0	0	0	0	0	0.0044	0	0	0	0.0039
PG	0	0	0	0	0	0	0	0	0	0	0.0105
RTX	0	0	0	0	0	0	0	0	0	0	0.0088
TRV	0	0	0	0	0	0	0	0	0	0	0.0053
UNH	0.6511	0.6578	0.3276	0.2329	0.1678	0.1682	0.0905	0.0907	0.0988	0.1235	0.1924
V	0	0	0.3276	0.2329	0.1678	0.1682	0.0905	0.0907	0.0988	0.1235	0.1908
VZ	0	0	0	0	0	0	0	0	0	0	0.0116
WBA	0	0	0	0	0	0	0	0	0	0	0.0063
WMT	0	0	0	0	0	0	0	0	0	0	0.0065
XOM	0	0	0	0	0	0	0	0	0	0	0.0084

Table 2.6: Portfolios Allocations for NDX 86 ($w = 0.02$)

i \ ε	**0 (Benchmark)**	0.0003	0.0007	0.001	0.004	0.01	0.018	0.03	0.05	0.1
AAPL	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
ADBE	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
ADI	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
ADP	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
ALXN	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
AMAT	0	0	0	0	0	0	0	0	0	0.0128
AMGN	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
AMZN	0	0	0	0	0	0	0	0	0.0154	0.0128
ANSS	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
ASML	0	0	0	0	0	0	0	0	0	0.0128
ATVI	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
BIIB	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
BMRN	0	0	0	0	0	0	0	0	0.0154	0.0128
CDNS	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
CERN	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
CHKP	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
CMCSA	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
COST	0	0	0	0	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
CPRT	0	0	0	0	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
CSCO	0	0	0	0	0	0	0	0	0.0154	0.0128
CSGP	0	0	0	0	0	0	0	0	0.0154	0.0128
CSX	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
CTAS	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
CTSH	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
CTXS	0	0	0	0	0	0	0	0	0	0.0128
DLTR	0	0.4000	0.2787	0.1702	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
DXCM	0.2709	0.2001	0.2787	0.1702	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
EA	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
EBAY	0	0	0	0	0	0	0	0	0	0.0128
EXC	0	0	0.1639	0.1492	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
EXPE	0	0	0	0	0	0	0	0	0	0.0128
FAST	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
FISV	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
GILD	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
GOOGL	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
IDXX	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
ILMN	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
INCY	0	0	0	0	0	0	0	0	0	0.0037
INTC	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0003	0.0007	0.001	0.004	0.01	0.018	0.03	0.05	0.1
INTU	0	0	0	0	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
ISRG	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
KLAC	0	0	0	0	0	0	0	0	0	0.0128
LBTYA	0	0	0	0	0	0	0	0	0.0154	0.0128
LBTYK	0	0	0	0	0	0	0	0	0.0154	0.0128
LRCX	0	0	0	0	0	0	0	0	0.0154	0.0128
LULU	0	0	0	0	0	0	0	0	0.0154	0.0128
MAR	0	0	0	0	0	0	0	0	0.0154	0.0128
MCHP	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
MDLZ	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
MNST	0	0	0	0	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
MSFT	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
MU	0	0	0	0	0	0	0	0	0	0.0128
MXIM	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
NFLX	0	0	0	0	0	0.0209	0.0021	0.0183	0.0122	0.0128
NTAP	0	0	0	0	0	0	0	0	0	0.0128
NTES	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
NVDA	0	0	0	0	0	0	0	0	0	0.0128
ORLY	0	0	0	0.1701	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
PAYX	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
PCAR	0	0	0	0	0	0	0	0	0	0.0128
PEP	0	0	0	0	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
QCOM	0	0	0	0	0	0	0	0.0144	0.0154	0.0128
REGN	0	0	0	0	0	0	0	0	0	0.0128
ROST	0	0	0	0.1702	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
SBUX	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
SIRI	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
SNPS	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
SWKS	0	0	0	0	0	0	0	0	0.0154	0.0128
TCOM	0	0	0	0	0	0	0	0	0	0.0128
TMUS	0	0	0	0	0	0.0392	0.0256	0.0183	0.0154	0.0128
TTWO	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
TXN	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
ULTA	0	0	0	0	0.0203	0.0392	0.0256	0.0183	0.0154	0.0128
VRSN	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128
VRTX	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
WBA	0	0	0	0	0	0	0	0.0183	0.0154	0.0128
WDC	0	0	0	0	0	0	0	0	0	0.0128
XEL	0.7291	0.4000	0.2787	0.1702	0.0891	0.0392	0.0256	0.0183	0.0154	0.0128
XLNX	0	0	0	0	0	0	0.0256	0.0183	0.0154	0.0128

Observations:

i. For the benchmarks ($\varepsilon = 0$) and $w = 0$ for DJI 29 and NDX 86 respectively, all shares are allocated to the one single stock that has the lowest $\bar{\beta}^*$. Figure 2.1 and 2.2 provides the scatterplot of $\bar{\alpha}_i^*$ and $\bar{\beta}_i^*$ for DJI 29 and NDX 86, respectivley. Note that $\bar{\alpha}_i^*$ and $\bar{\beta}_i^*$ are computed by using Eqs. (2.11) and (2.12), which are derived from Ordinary Least Squares estimation.

   When compare the allocations the scatterplot in Figure 2.1 and 2.2, note that the order in which the computer selects stocks when it increases diversification is similar to a waterfilling procedure where we start with the stock with the lowest $\bar{\beta}^*$ and then select stocks with the next lowest $\bar{\beta}^*$ up until either full diversification is achieved or we reach a critical value of $\bar{\beta}^*$, before the number of stocks we invest in decreases.

ii. For a given value of w, when ε gradually increases, the diversification of the optimal allocation increases for DJI 29 and NDX 86.

   a. $w = 0$:

      The portfolio allocation is equally split into the stocks the manager invests in (those stocks i where $x_i > 0$ at optimality). In DJI 29 (see Table 2), when $\varepsilon = 0$, the entire portfolio is put into MCD, which has the lowest $\bar{\beta}^*$. MCD is shown in the red box in Figure 2.1. When ε increases, the allocation is equally split among stocks with small $\bar{\beta}^*$. The number of stocks invested in increases as ε increases. The same observation holds for the NDX 86 data set. (When $\varepsilon = 0$, all the allocation is put into EXC, which has the lowest $\bar{\beta}^*$. EXC is shown in Figure 2.2.)

   b. $w > 0$:

      For problems that have a nonempty feasible set for the selected w, there are values

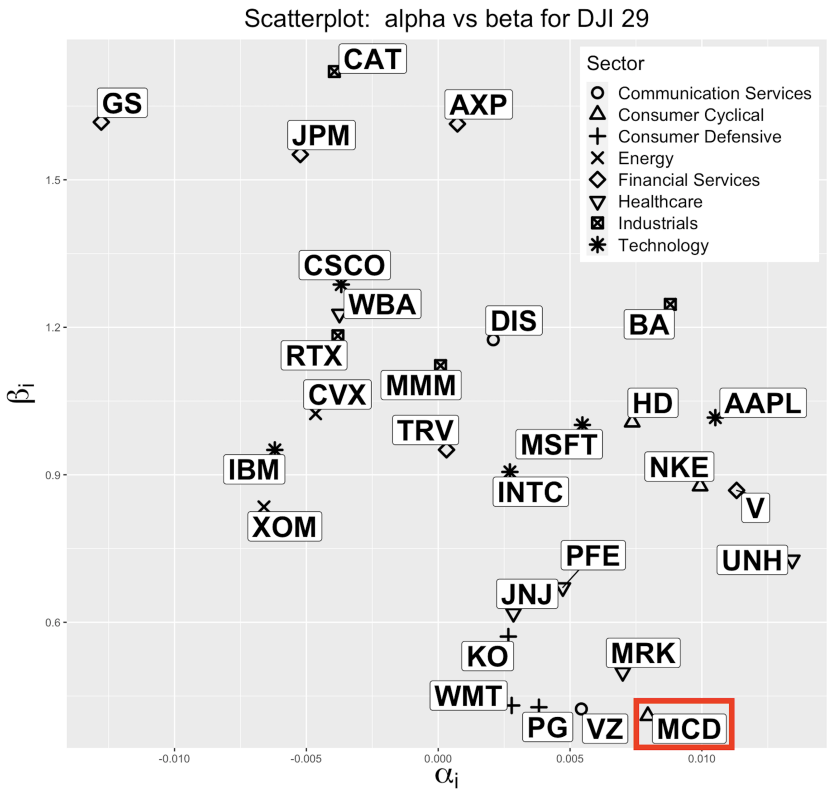

Figure 2.1: Scatterplot of $\bar{\alpha}_i^*$ and $\bar{\beta}_i^*$ for DJI 29

of w for which shares of allocated stocks are not all evenly divided; however, when it happens, all stocks invested in except one receive the same allocation (as fraction of the overall portfolio) and the last stock invested in receives a smaller allocation.

c. ε vs number of allocated stocks:

Figures 2.3 and 2.4 show the scatterplots of ε vs number of allocated stocks for DJI 29 and NDX 86 respectivley, where the dot is the number of allocated stocks for a given ε. Except for DJI 29 at $w = 0.015$ and DJI 29 at $w = 0.02$, the number of stocks invested in increases when ε increases.

The graphs for DJI 29 at $w = 0.015$ and DJI 29 at $w = 0.02$ are similar to those obtained in the classical Bertsimas-Sim [19] robust optimization frameworks where the number of stocks invested in increases as the budget of uncertainty increases until a breakpoint value for the budget of uncertainty is reached, and then decreases as the budget of uncertainty increases past the breakpoint. This happens

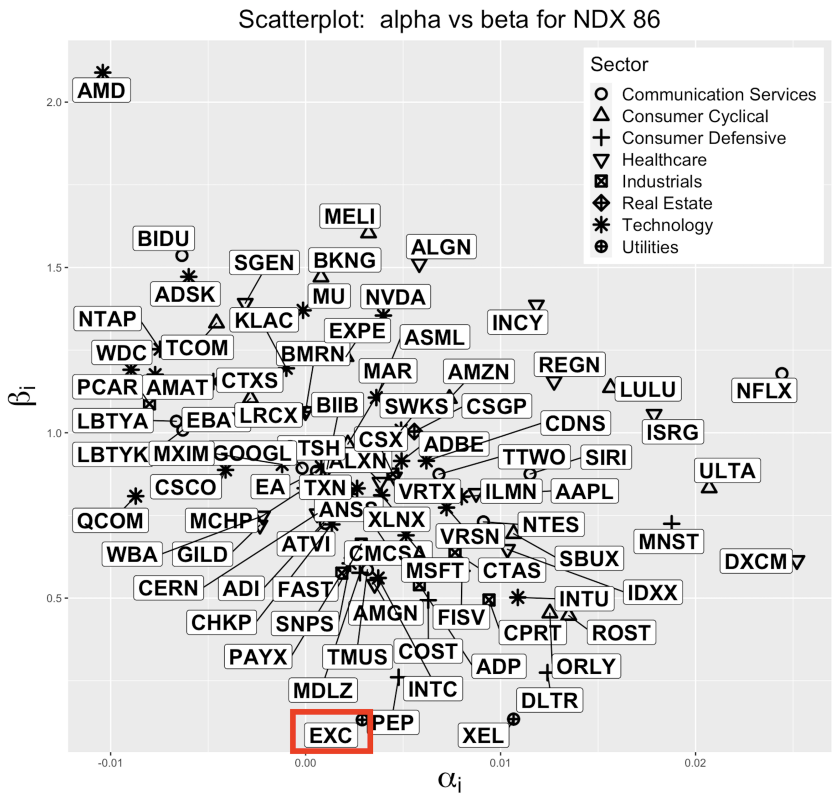

Figure 2.2: Scatterplots of $\bar{\alpha}_i^*$ and $\bar{\beta}_i^*$ for NDX 86

when an increasingly pessimistic outlook from the manager leads to invest in the "safer" stocks instead of diversifying. Note that the robust problems corresponding to DJI 29 for $w = 0.015$ and DJI 29 for $w = 0.02$ are infeasible when $\varepsilon > 0.07$ and $\varepsilon > 0.023$, respectively.

2.4.2. Numerical Experiments for Simulation

This section tests the performance of portfolio in a real-life environment where simulating r_{mt}, and r_{it} for all i and t by introducing randomness on the actual $\bar{\alpha}_i^*$, $\bar{\beta}_i^*$ and the residual returns e_{it} with $T = 10,000$ scenarios for both DJI 29 and NDX 86. Consider simulation experiments for $w = 0$, $w = 0.01$, $w = 0.015$ and $w = 0.02$ varying values of ε, and only $\varepsilon's$ that substantially change the portfolios structures $x_i's$ are recorded. The expected return $\sum_i^N r_{it}x_i \ \forall \ t = 1 \dots T$ and volatility between the benchmark ($\varepsilon = 0$) and the portfolios

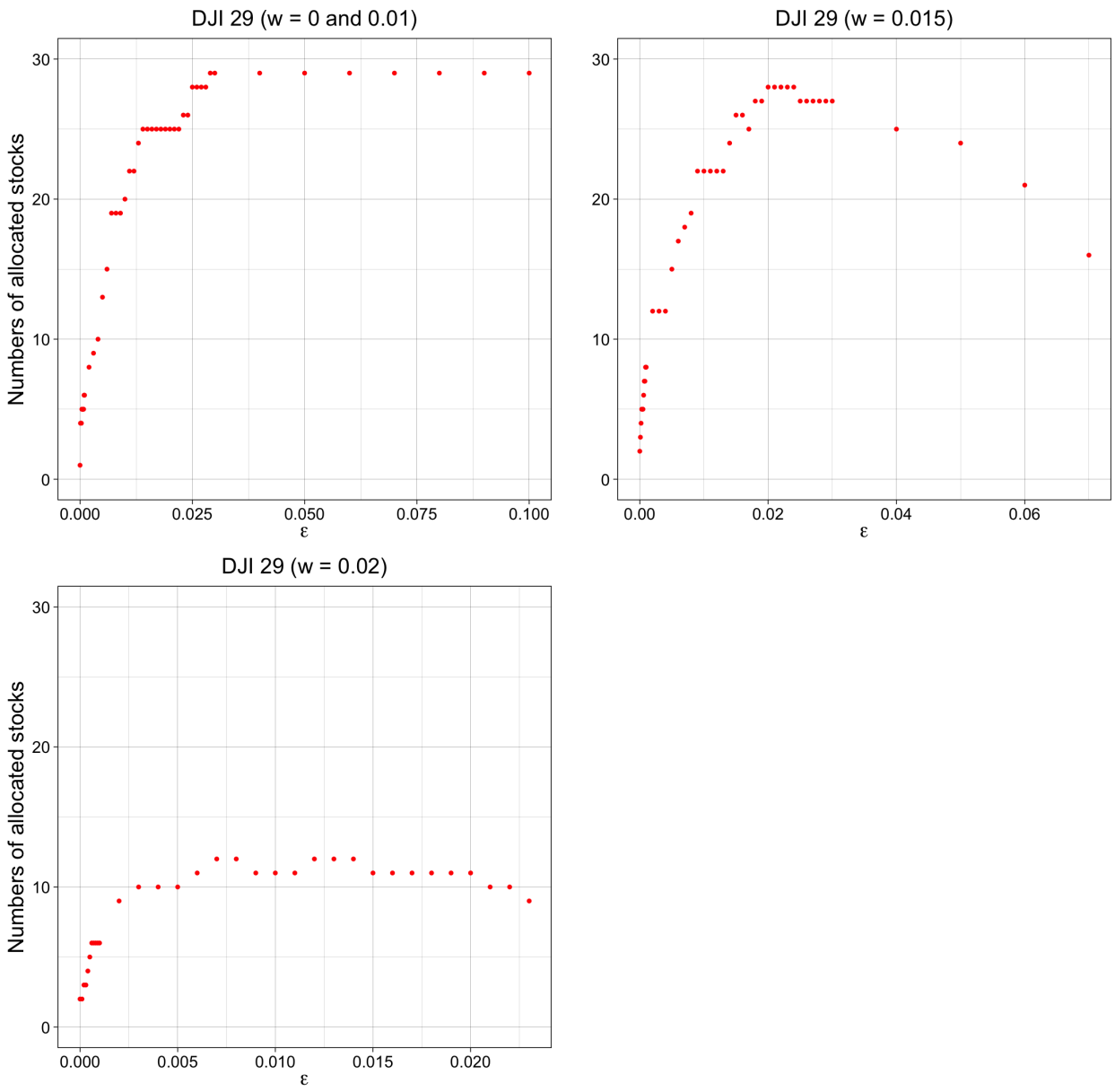

Figure 2.3: ε vs. Numbers of Allocated Stocks for DJI 29

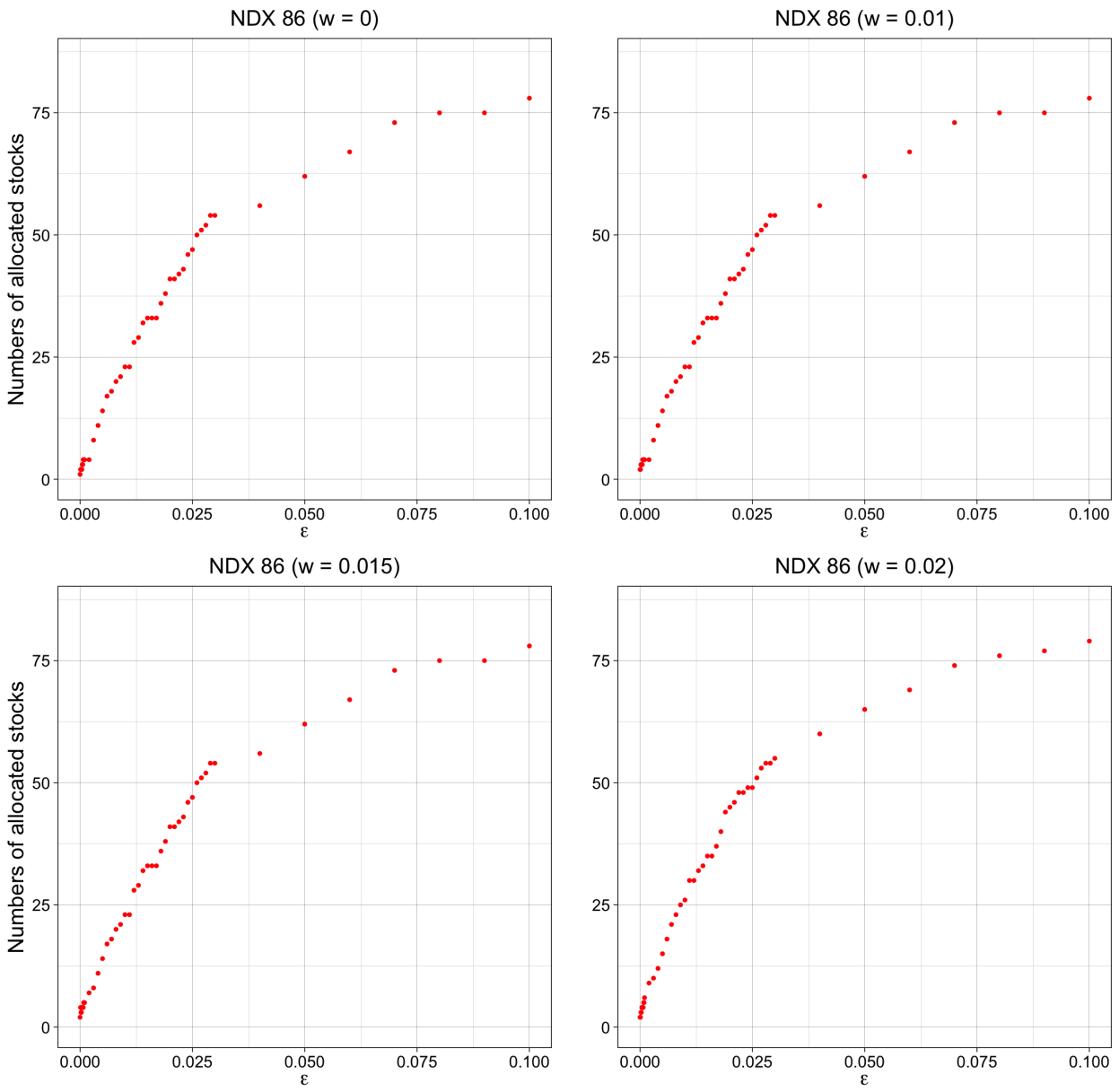

Figure 2.4: ε vs. Numbers of Allocated Stocks for NDX 86

obtained for different ε are compared, for a given level of w. The robust solutions $x_i's$ for each w are computed by solving the proposed model **RP**.

2.4.2.1. Procedures for Simulation Experiments

The following steps are followed to perform the DJI 29 and NDX 86 simulation experiments:

Step 1. Simulating the return on the market index $\tilde{r}_{mt}$: We compute the average return on the market index $\hat{\mu}_m$ and standard deviation σ_m from the historical monthly data sets (March 1st, 2009 - February 1st, 2019). We generate $T = 10,000$ data points from a Normal distribution with mean $\hat{\mu}_m$ and standard deviation σ_m to obtain simulated return on the index denoted by $\tilde{r}_{mt}$.

Step 2. Simulating the return on stock $\tilde{r}_{it}$: In Sharpe's single-index model, the return of stock i in time period t is modeled as $r_{it} = \alpha_i + \beta_i r_{mt} + e_{it}$ (2). We simulate $\tilde{\alpha}_i$, $\tilde{\beta}_i$ and residual returns $\tilde{e}_{it}$.

i. Simulate $\tilde{\alpha}_i$ and $\tilde{\beta}_i$:

We add a random term drawn from a standard normal distribution with mean 0 and standard deviation 1 scaled by $\pm 20\%$ of the estimated $\bar{\alpha}_i$ and $\bar{\beta}_i$ to the nominal parameters in order to simulate the actual, unknown $\tilde{\alpha}_i$ and $\tilde{\beta}_i$, respectively.

ii. Simulate residual returns $\tilde{e}_{it}$:

We simulate $10,000$ residual returns denoted by $\tilde{e}_{it}$ that are Normally distributed with mean $\hat{\mu}_{ei} = \frac{1}{T}\sum_{t=1}^{T} e_{it}$, and standard deviation σ_{ei} $\forall i \in N$.

iii. Generate simulated returns $\tilde{r}_{it}$:

$\tilde{r}_{it} = \tilde{\alpha}_i + \tilde{\beta}_i \tilde{r}_{mt} + \tilde{e}_{it} \qquad \forall i \in N,\ \forall t \in T.$

Step 3. Compute $\sum_i^N \tilde{r}_{it} x_i \ \forall \ t = 1 \dots T$, and draw density plots and boxplots for each combination of w and ε to compare the performance of our robust solutions with the performance of benchmarks.

2.4.2.2. Results for Simulation Experiments

In Figures 2.5 and 2.7 (DJI 29), density plots and boxplots are included to compare the performance of the robust model to that of the benchmark in 10,000 simulated instances.

For both simulated DJI 29 and NDX 86, the red curves in density plots and boxplots in the first columns represent the benchmarks. In addition, the solid dots inside the boxplots represent the mean returns of portfolios $\frac{1}{T} \sum_t^T \sum_i^N \tilde{r}_{it} x_i \ \forall w, \varepsilon$. The tables after the figures show the mean, median and standard deviation of the portfolio returns $\sum_i^N \tilde{r}_{it} x_i$ for $t = 1 \dots T$. To make the data in the tables easier to read, lower and higher values are highlighted in gradient lighter and darker green, respectively.

In the simulated DJI 29, for each value of w, the mean returns and the medians (black solid lines in boxplots) with $\varepsilon > 0$ are slightly higher than the benchmark ($\varepsilon = 0$), or very close to the benchmark. Furthermore, there is an optimal $\varepsilon > 0$ that achieves a lower volatility. For $w = 0$ and $w = 0.01$ (see Figure 2.5), when ε is between 0.0001 and 0.002, the interquartile ranges (IQR) of the boxplots are much narrower. The same observation holds for $w = 0.015$ (see appendix) and $w = 0.02$ (see Figure 2.7) when ε is between 0.0005 and 0.001; however, when ε is out of this range, the advantage provided by the lower volatility of our robust solutions over the benchmark shrinks.

Decision makers prefer ε that lead to portfolios with higher mean and median, and a lower standard deviation. There may exist multiple such ε. First, find candidate ε's that outperform the benchmark in terms of both higher mean and median, and lower standard deviation, and then remove from the candidate set values of ε that are dominated by other

candidate values of ε (i.e., they lead to portfolio returns that have lower mean and median, and higher standard deviation.) Decision makers can then choose an ε from the remaining candidates based on their risk preferences.

For DJI 29 ($w = 0$ and 0.01) in Table 2.7, all $\varepsilon > 0$ outperform the benchmark; however, the portfolio for $\varepsilon = 0.0001$ has lower mean and median, and a higher standard deviation compared to $\varepsilon = 0.0008$, 0.009 and 0.002. Thus, $\varepsilon = 0.0001$ is removed from the candidate set. Similarly, $\varepsilon = 0.0008$ and 0.0009 are outperformed by $\varepsilon = 0.002$. Moreover, $\varepsilon = 0.002$ is not dominated by the remaining ε since it has lower standard deviation. Therefore, $\varepsilon = 0.002$, 0.007, 0.011, 0.023 and 0.029 are in the candidate set for DJI 29 ($w = 0$ and 0.01). For DJI 29 ($w = 0.02$) in Table 2.9, $\varepsilon = 0.0005$, 0.0009, 0.001, 0.014 and 0.02 are in the candidate set.

Table 2.7: Simulated DJI 29 ($w = 0$ and 0.01)

ε	0	0.0001	0.0008	0.0009	0.002	0.007	0.011	0.023	0.029
Mean	0.011	0.011	0.012	0.012	0.013	0.015	0.014	0.015	0.015
Median	0.011	0.011	0.012	0.012	0.013	0.015	0.015	0.015	0.016
S.D.	0.039	0.027	0.026	0.025	0.025	0.029	0.030	0.032	0.036

Table 2.8: Simulated DJI 29 ($w = 0.015$)

ε	0	0.0001	0.0005	0.0006	0.0008	0.001	0.005	0.014	0.03	0.07
Mean	0.013	0.015	0.015	0.016	0.016	0.016	0.015	0.015	0.016	0.020
Median	0.013	0.015	0.015	0.016	0.016	0.016	0.015	0.016	0.016	0.020
S.D	0.037	0.033	0.030	0.028	0.028	0.027	0.027	0.032	0.036	0.041

Table 2.9: Simulated DJI 29 ($w = 0.02$)

ε	0	0.0001	0.0004	0.0005	0.0009	0.001	0.008	0.014	0.02	0.023
Mean	0.020	0.020	0.020	0.020	0.019	0.019	0.020	0.021	0.022	0.022
Median	0.019	0.019	0.020	0.020	0.019	0.019	0.020	0.021	0.022	0.022
S.D	0.044	0.044	0.037	0.034	0.033	0.033	0.038	0.042	0.044	0.045

In simulated NDX 86, the advantage of robust solutions over the benchmark is magnified.

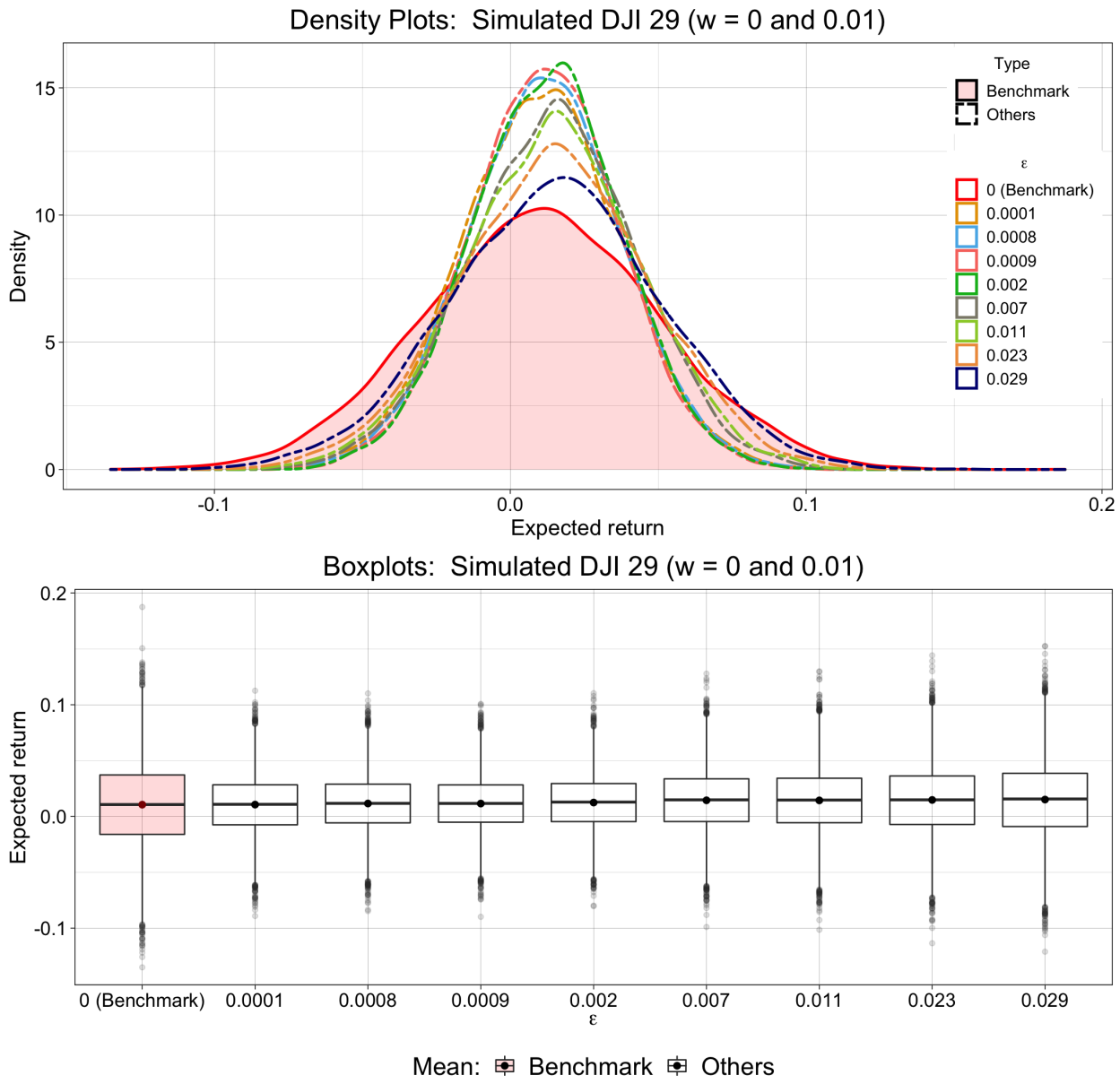

Figure 2.5: Performance on Simulated DJI 29 ($w = 0$ and 0.01)

Table 2.10: Simulated NDX 86 ($w = 0$)

ε	0	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
Mean	0.007	0.009	0.010	0.012	0.016	0.018	0.019	0.019	0.020	0.020	0.021
Median	0.007	0.009	0.010	0.012	0.016	0.018	0.019	0.020	0.020	0.020	0.021
S.D	0.053	0.035	0.027	0.026	0.025	0.028	0.032	0.034	0.036	0.037	0.041

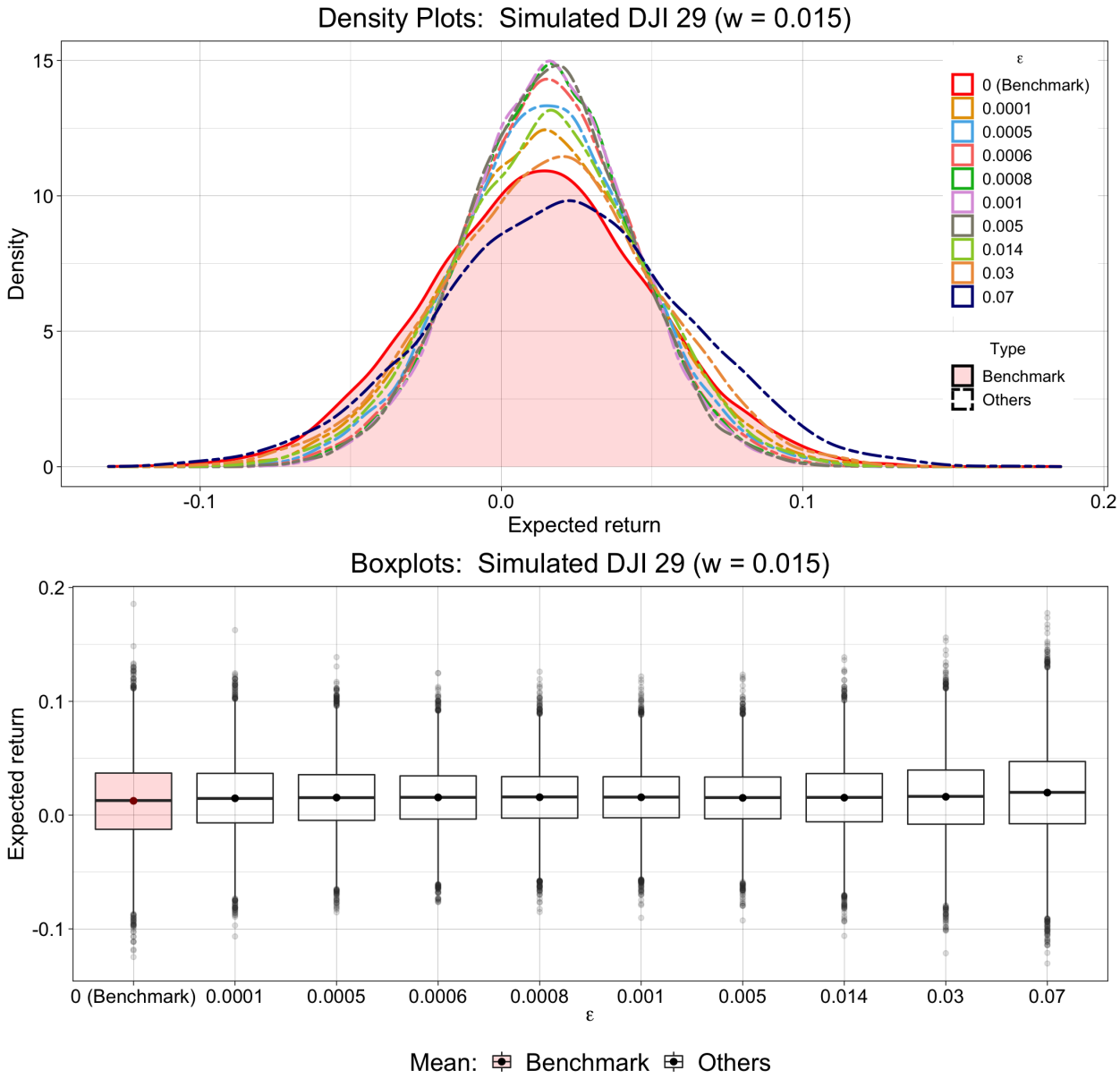

Figure 2.6: Performance on Simulated DJI 29 ($w = 0.015$)

Table 2.11: Simulated NDX 86 ($w = 0.02$)

ε	0	0.0003	0.0007	0.001	0.004	0.01	0.018	0.03	0.05	0.1
Mean	0.020	0.021	0.022	0.021	0.021	0.021	0.020	0.020	0.020	0.021
Median	0.020	0.021	0.022	0.021	0.021	0.021	0.020	0.020	0.020	0.021
S.D	0.048	0.042	0.044	0.034	0.032	0.031	0.033	0.035	0.037	0.041

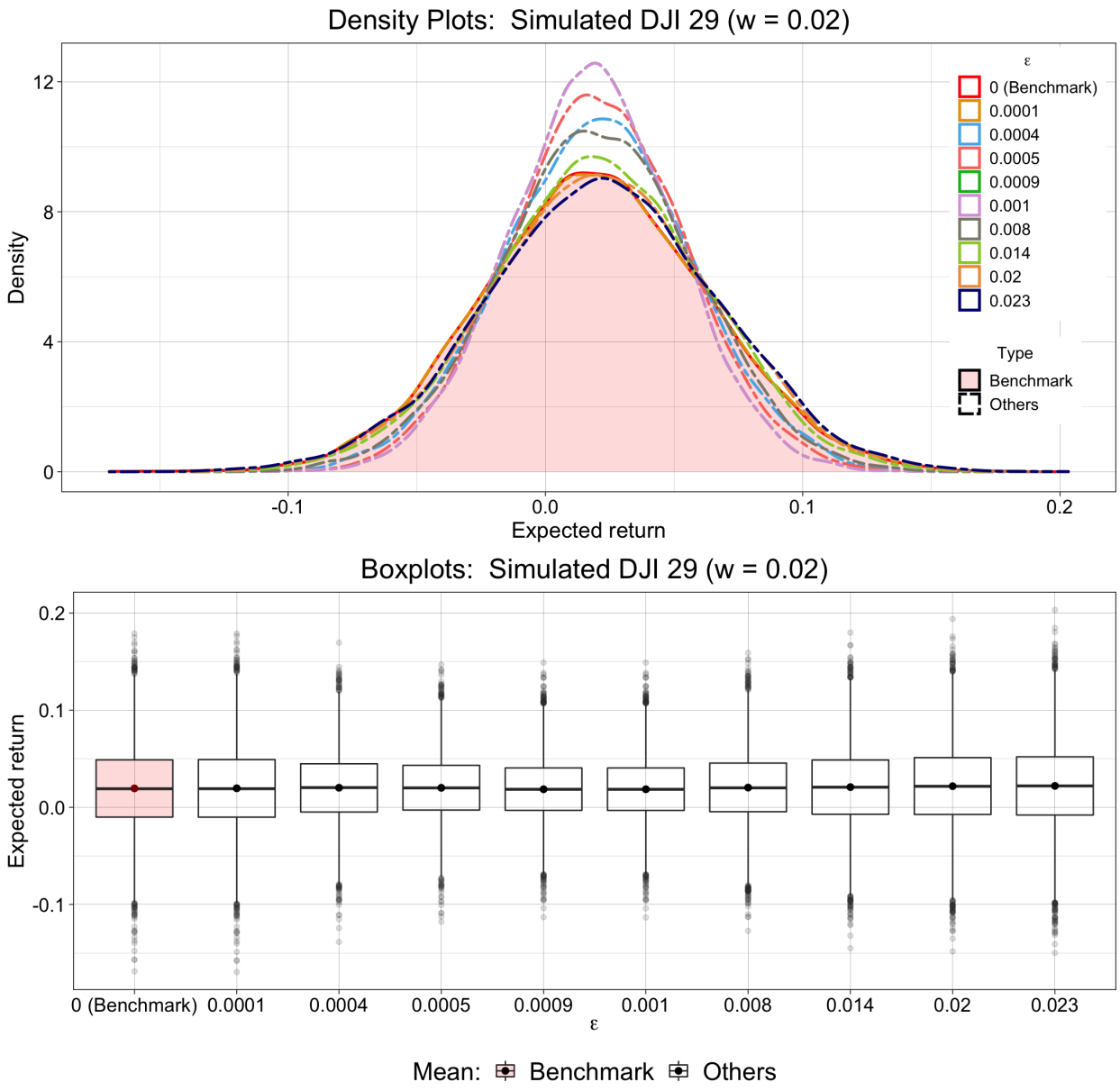

Figure 2.7: Performance on Simulated DJI 29 ($w = 0.02$)

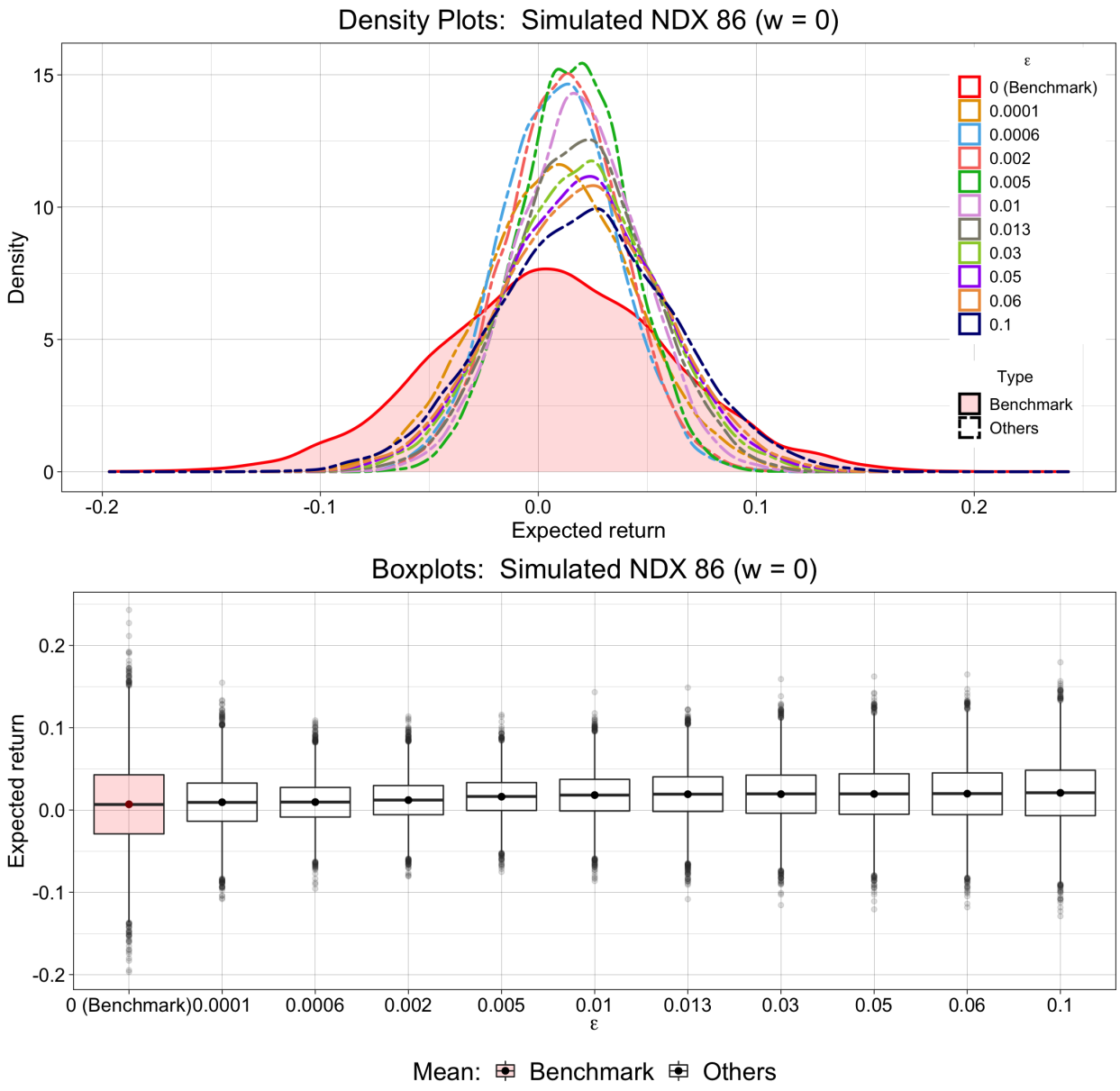

Figure 2.8: Performance on Simulated NDX 86 ($w = 0$)

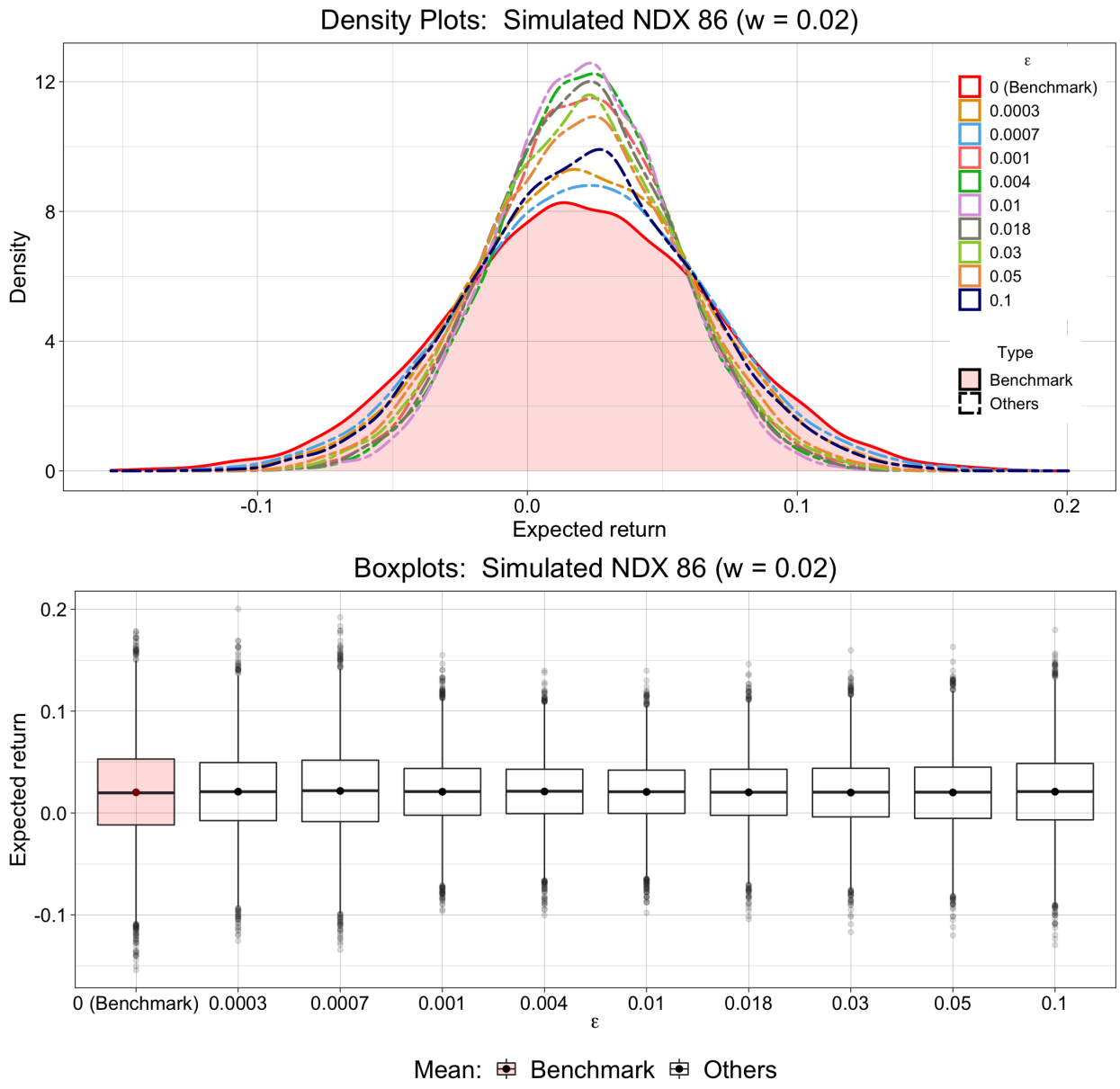

Figure 2.9: Performance on Simulated NDX 86 ($w = 0.02$)

2.4.2.3. Comparison with Goldfarb and Iyengar's Model

In this section we compare the performance of our robust portfolio with the robust minimum variance problem in [18]. Their model includes a prespecified confidence threshold ω. In their words, "If ω is chosen very high, the uncertainty sets will be very large. On the other hand, if ω is chosen too low, the portfolio choice will not be robust enough. The typical choices of ω lie in the range $0.95 - 0.99$." ([18] p. 18). $\omega = 0.95$ is set in this comparison.

In Goldfarb and Iyengar's model, the problem is infeasible when $w = 0.015$ and $w = 0.02$. Therefore, the numerical study below focuses on $w = 0$ and $w = 0.01$.

For DJI 29, the allocations for $w = 0$ and $w = 0.01$ lead to the same boxplots, shown in Figure 2.10. We observe that our portfolios offer lower volatility except for $\varepsilon = 0$ and $\varepsilon = 0.029$, and the portfolios for $\varepsilon = 0$ and $\varepsilon = 0.029$ exhibit higher upside risk. For $\varepsilon = 0.007$, 0.011 and 0.023, the proposed **RP** outperforms Goldfarb and Iyengar's model ($w = 0$ and 0.01) with respect to the mean, median and standard deviation of the portfolio return.

Table 2.12: Comparison with Goldfarb and Iyengar's model for DJI 29 ($w = 0$ and 0.01)

ε	0	0.0001	0.0008	0.0009	0.002	0.007	0.011	0.023	0.029	Goldfarb($w = 0$)	Goldfarb ($w = 0.01$)
Mean	0.011	0.011	0.012	0.012	0.012	0.014	0.014	0.015	0.015	0.010	0.014
Median	0.011	0.011	0.012	0.012	0.013	0.015	0.015	0.015	0.016	0.010	0.015
S.D	0.039	0.027	0.026	0.025	0.025	0.028	0.030	0.032	0.036	0.033	0.034

Table 2.13: Comparison with Goldfarb and Iyengar's model for NDX 86 ($w = 0$)

ε	0	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1	Goldfarb($w = 0$)
Mean	0.007	0.009	0.010	0.012	0.016	0.018	0.019	0.019	0.020	0.020	0.021	0.024
Median	0.007	0.009	0.010	0.012	0.016	0.018	0.019	0.020	0.020	0.020	0.021	0.024
S.D	0.053	0.035	0.027	0.026	0.025	0.028	0.032	0.034	0.036	0.037	0.041	0.039

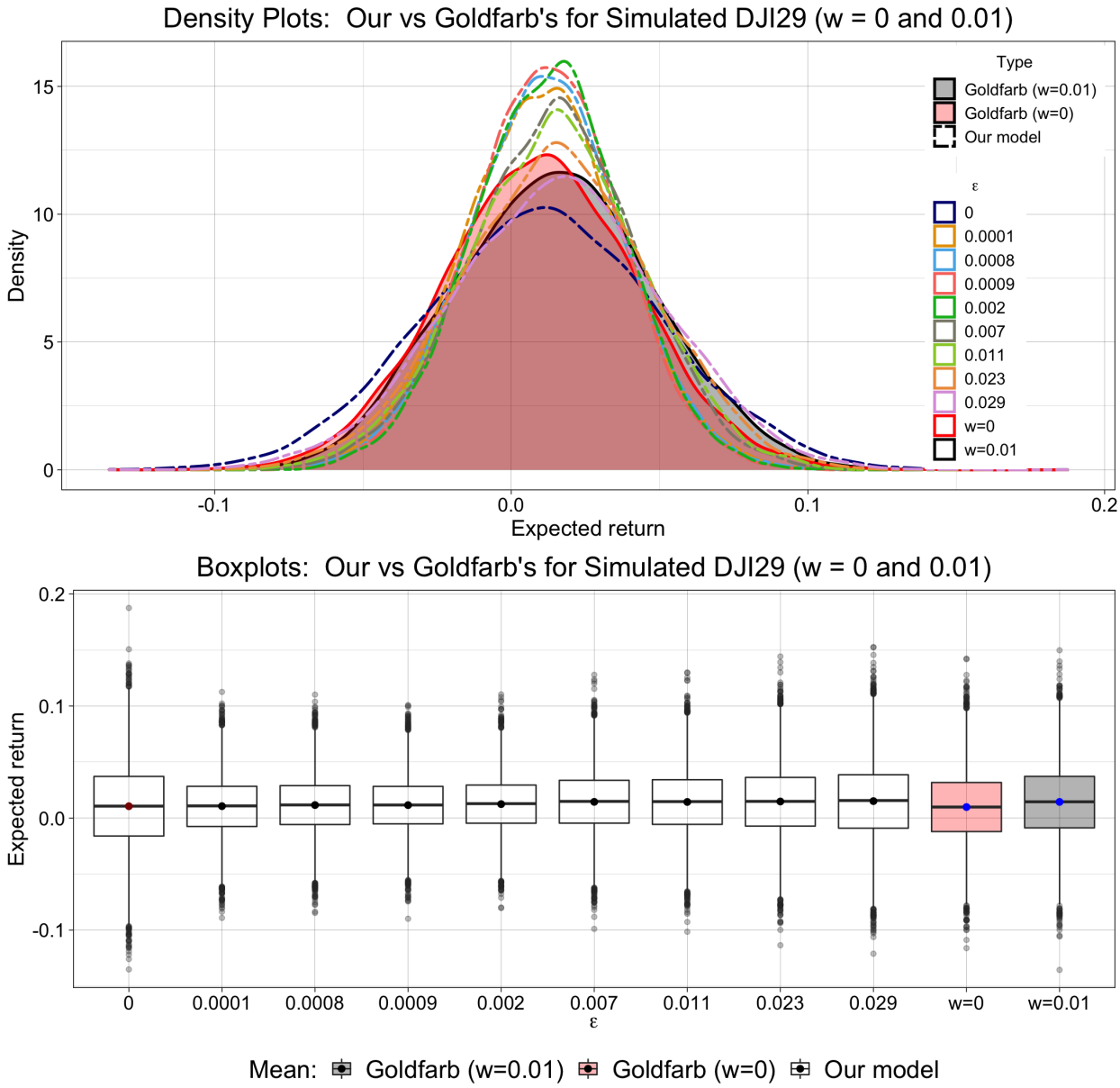

Figure 2.10: Comparison with Goldfarb and Iyengar's model ($w = 0$ and $w = 0.01$ for DJI 29)

Table 2.14: Comparison with Goldfarb and Iyengar's model for NDX 86 ($w = 0.01$)

ε	0	0.0005	0.002	0.005	0.011	0.017	0.02	0.04	0.06	0.1	Goldfarb($w = 0.01$)
Mean	0.010	0.010	0.012	0.016	0.018	0.019	0.019	0.019	0.020	0.021	0.024
Median	0.009	0.010	0.012	0.016	0.018	0.019	0.019	0.020	0.020	0.021	0.024
S.D	0.034	0.032	0.026	0.025	0.028	0.032	0.032	0.034	0.037	0.041	0.042

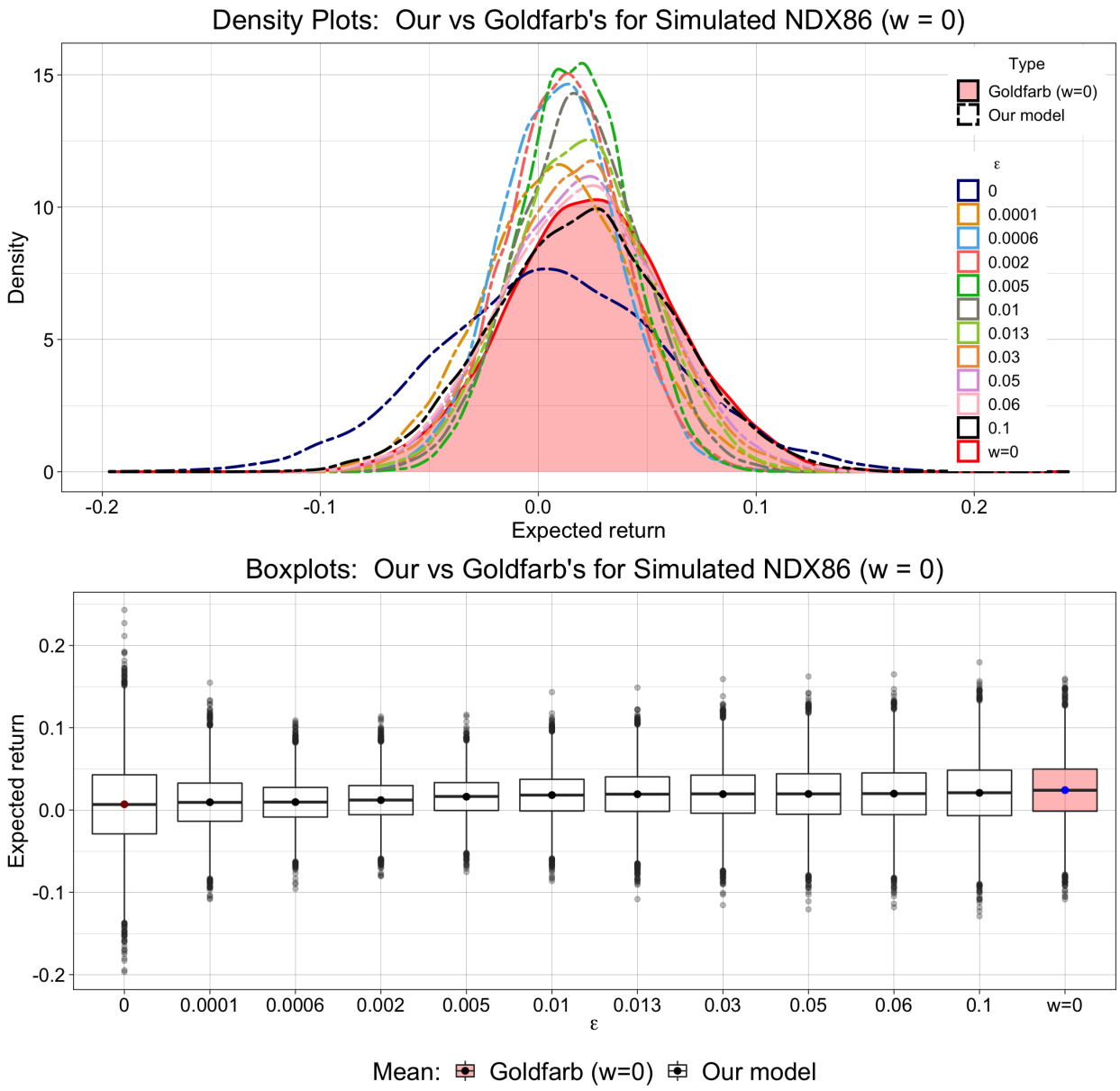

Figure 2.11: Comparison with Goldfarb and Iyengar's model ($w = 0$ for NDX 86)

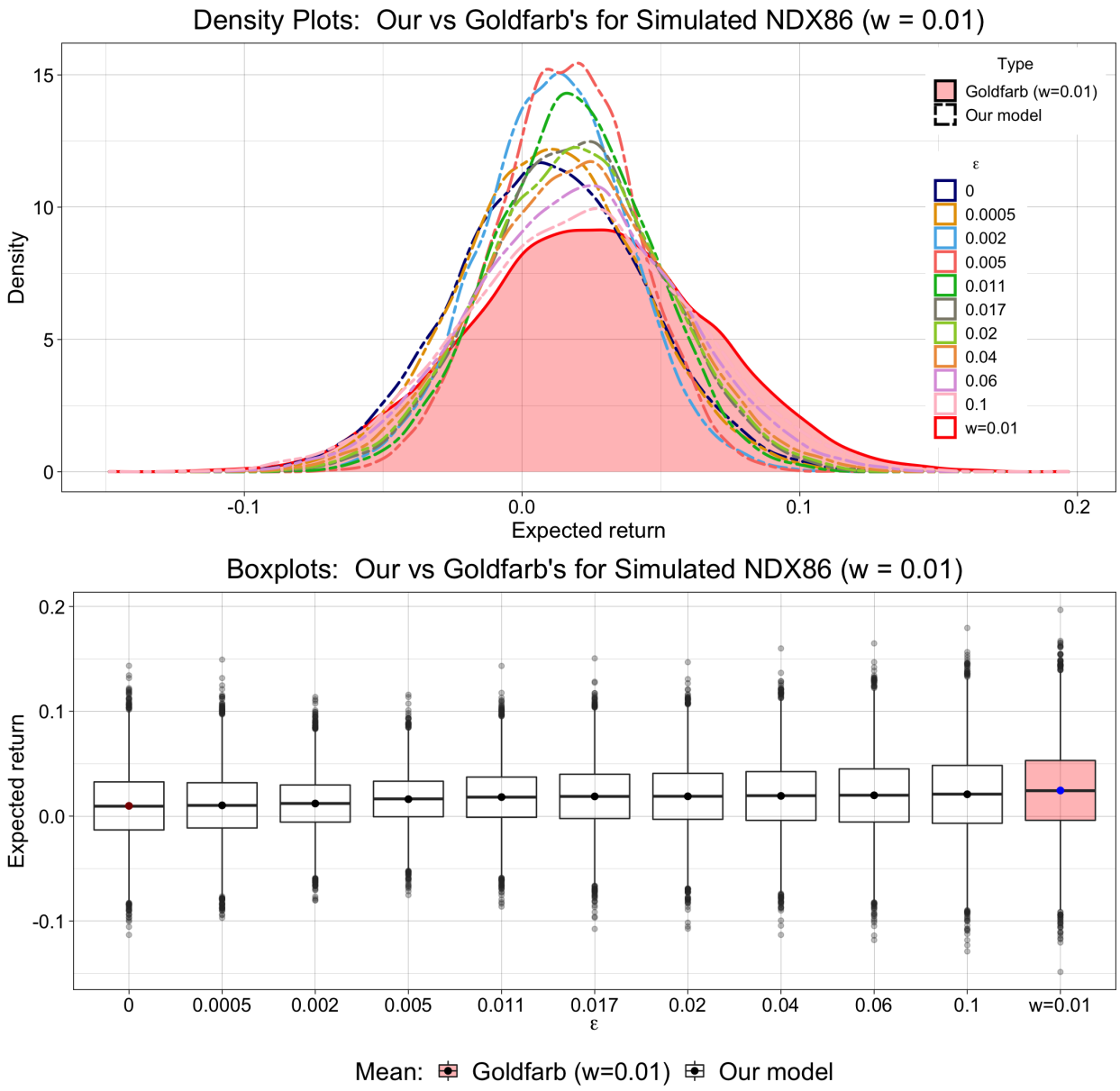

Figure 2.12: Comparison with Goldfarb and Iyengar's model ($w = 0.01$ for NDX 86)

The key advantage of the proposed approach is that it can generate portfolios with lower volatility regarding standard deviation. For NDX 86 with $w = 0$ and $w = 0.01$, allocation of Goldfarb and Iyengar's model provides both higher means, medians, and standard deviations. However, their approach does not provide the decision-makers with an option to decrease the risk based on their risk preference. Hence, the advantage of the proposed approach is to provide the decision maker with a menu of portfolio allocations from which he can choose to better match his risk appetite.

2.5. Summary

In this study, a robust optimization model for Mean-Variance portfolio selection has been proposed. The model considers Markowitz's Mean-Variance Optimization when stock returns follow Sharpe's single-index model. However, in real-world settings, the Sharpe model coefficients are not always precisely known. Therefore, a tractable robust optimization model has been developed, assuming that coefficients α and β, estimated using least-square estimators, are within a prespecified distance of optimality.

Numerical experiments were conducted, in which the minimum required expected return from the portfolio was varied and the performance of the model was analyzed when the allocation structure changed significantly. In practical applications, it would be useful to consider a wider range of values for w and ε with smaller step sizes. The decision maker can choose epsilon based on their preference for the trade-off between expected return and volatility.

We have conducted numerical experiments on both Dow Jones and NASDAQ 100 indices. The results indicate the existence of an optimal $\varepsilon > 0$, where the portfolio realizes a slightly higher return with lower volatility in terms of mean and standard deviations, respectively. Furthermore, when the value of ε increases, an increasing number of stocks will be allocated

evenly in most cases. In future research, we intend to extend our robust model to a standard multiple-index portfolio selection problems (e.g., the Fama-French Model).

Chapter 3

Renewable Energy

3.1. Introduction

In recent decades, the development and implementation of renewable energy have gained increasing attention. The utilization of renewable energy has become more prominent due to its numerous positive impacts on the climate and public health [27, 28]. Several studies have evaluated the future potential and utilization of renewable resources. Ellabban et al. [29] assessed the availability and potential of renewable energy sources globally and ample potential for solar, wind, geothermal, and biomass energy. Lior [30] analyzed projections for energy demand that could be met by renewable sources and studied potential benefits and limitations of the environment and economy. Hadjerioua et al. [31] conducted a comprehensive assessment of hydroelectric potentials in the United States. Their study quantifies the potential capacity for non-powered dams and highlighted areas with significant unexploited resources. Brooks [32] provided comprehensive evaluations of renewable potential resources and current installations in the United States. Lopez et al. [33] provided both national and state-level estimation of technical potential capacity and geographical limitation for renewable energy.

To promote the use of renewable technologies and reduce dependence on fossil energy by requiring a certain percentage of the electricity generated from renewable sources, a non-mandatory Renewable Portfolio Standards (RPS) policy was implemented in 30 states and Washington, D.C. in the United States as of 2021 [34]. Nogee et al. [35] reviewed an analysis of the history and projected impacts of national RPS proposals. Wiser et al. [36] outlined a roadmap and summarized the cost and benefit of state-level RPS. Regression

analysis is commonly applied in the literature to evaluate the non-projected effectiveness of RPS [37–40]. Huang et al. [37]. studied factors such as education level and political party affect the adoption of RPS by states. Yin and Powers [38] applied regression analysis to study the increase in renewable generation resulting from state RPS. Fischlein and Smith [39] also used regression analysis on the outcome and assessment of state-by-state RPS policy design. Recently, Joshi [40] studied the expansion of the implementation of RPS across states. The literature that utilizes regression analysis is more oriented toward the retrospective evaluation of RPS. Mai et al. [41–43] provided a relatively comprehensive study on the future influences of state-level RPS by adopting a deterministic linear program while converting pollutant emissions, public health, and economic growth into monetary values.

For the optimization modeling, Ding and Somani [44] propose a deterministic linear programming model for a single-period wind power investment planning that meets the RPS target for the Midwest Independent System Operator. Lara et al. [45] develop a deterministic Mixed Integer Linear Programming (MILP) model for power planning and apply it to the Electric Reliability Council of Texas (ERCOT). They propose their own decomposition algorithm to solve this intractable problem. Building on Lara et al.'s model, Li et al. [46] introduce simplifications to the original problem and a tailored Benders decomposition algorithm.

In real-world energy systems, energy planning integrated with renewable energy faces uncertainty. These uncertainties include variations in electricity demand driven by changes in population and economic activities. Additionally, fluctuations in energy prices and operation and maintenance costs further complicate energy planning. To address the challenge, it is critical to capture the uncertainty in an energy planning model. Robust optimization (RO) is a method for addressing uncertainties in energy planning with renewable sources [47–49]. Mulvey et al. [47] examine the advantage of robust optimization in a large-scale system. They apply the robust optimization technique to a simple single-period plant capacity expansion

problem, where the goal is to minimize the capacity expansion cost while satisfying the demand in a single period. Koo et al. [48] incorporate carbon dioxide requirement and emissions trading in their formulation. Akbari et al. [49] apply robust optimization to the investment in a distributed energy system serving a small geographic area. Rezvan et al. [50] proposed a RO approach to finding the optimal capacity of generation systems to optimize multiple objectives under the uncertainty in energy demand. Parisio et al. [51] analyzed a robust operation scheduling problem for multi-generation systems, in the face of uncertain energy costs and demands. Lorca and Sun [52], and Zugno and Conejo [53] used RO apporach to minimize the system costs in terms of day-ahead energy dispatch. Xiong and Singh [54] applied two-stage distributionally RO for the energy and reserve under uncertain RE power. Yu et al. [55] used RO technique to determine the operation of the large electricity consumers problem that considered RE sources under the uncertain market price. RO approach is also commonly applied in the field of renewable distributed generation, which involves the implementation of small-scale RE production systems [49,56–58]. Moret et al. [59] extended a RO framework for a single-period energy planning that accounted for multiplied uncertainties in the objective.

While there has been a growing trend in recent years, the implementation of RO approach in long-term renewable energy planning is still limited. To the best of my knowledge, the existing literature does not consider the optimization of long-term renewable energy planning state-level RPS targets with uncertainties.

Contributions:

In this study, we propose a comprehensive long-term energy planning model that integrates renewable technologies to achieve environmental targets. Unlike typical single-period planning models in the literature that assume capacity expansion is immediately available at any time, our proposed model incorporates construction lead times required before new infrastructure is operational. We derive a closed-form expression for installed capacity over

the planning horizon that captures the construction's lead times. In addition, our model considers the limited resources for capacity expansion in the market that imposes additional costs on the competition. In other works in the literature, the cost is the same for any capacity expansion. Furthermore, we develop a robust optimization model for uncertain resource cost rates, variable and fixed operation and maintenance cost rates, electricity purchase prices, and electricity demand. Robust optimization is well-suited for this purpose due to the imprecise information available on rates, prices and demands for future years. For the demand, we provide decision-makers with options to control the uncertainty in each time period. We provide a tractable robust formulation that can be solved by a commercial off-the-shelf solver. In the numerical experiments, we apply our approach to California and provide decision-makers with various strategies for capacity expansion and electricity generation profiles. To the best of our knowledge, ours is the first work in energy planning that considers a decades-long planning horizon, construction lead times needed to make new energy infrastructure operational, different costs for different expansions, and implement robust optimization; we also provide a comprehensive approach to estimate the parameters needed to implement the model using real data.

The rest of the chapter is structured as follows. In Section 3.2, we propose a deterministic formulation for the energy planning model. In Section 3.3, we construct uncertainty sets, develop a robust approach, and provide a tractable formulation. We conduct numerical experiments in Section 3.4.

3.2. Deterministic Model

In this section, we introduce a deterministic long-term energy planning model integrated with renewable energy. The problem can be formulated as the decision makers aim to plan to invest in energy infrastructure and design electricity generation profiles to meet future demands and achieve planned environmental goals while minimizing the total costs over decades. The problem can be briefly formulated as follows, and the full mathematical model is introduced in section 3.2.2:

- Objective: Minimize the total costs incurred in all time periods.

  The total costs include:

  1. Infrastructure investment costs.
  2. Recourse or fuel costs for electricity generation.
  3. Variable operation and maintenance (O&M) costs.
  4. Fixed O&M costs.
  5. Electricity purchase costs

- Subject to installed capacity representation:

  The installed capacity at a time period t is decided by the capacity in the time period $t-1$, capacity expansion, planned capacity addition, planned capacity retirement, and unplanned retirement

- Electricity generation constraint:

  The generation is restricted by several factors, including the capacity factor of generating technologies, the installed capacity, and transmission loss.

- Electric reliability constraint:

  To ensure electrical reliability and avoid extreme situations, additional capacity is required. In our problem, these reserve margin is covered by conventional technologies.

- Potential capacity constraint:

  Due to the geographical attributes, the installed capacity for renewable technologies is limited by potential capacity

- Demand constraint:

  The demand is satisfied by electricity generation and purchase.

- Environmental targets constraints:

  1. Renewable Portfolio Standards (RPS) require a certain percentage of electricity to be generated from renewable technologies.
  2. Clean Electricity requires a certain percentage of electricity to be generated from technologies with zero emissions.
  3. A certain percentage of electricity is required to be generated from a specific renewable technology.
  4. A required amount of electricity is generated by renewable technologies.

- Purchase constraint:

  Electricity purchase is limited by an upper bound. In addition, the purchase growth rate is also restricted.

- Capacity expansion constraint:

  The capacity expansion by technology is restricted by an upper limit. Surpassing a particular threshold in capacity expansion will lead to a surge in costs due to resource competition in the market.

3.2.1. Nomenclature

We use the following notations throughout this chapter:

Sets

T: the set of time periods of energy planning,

I: the set of electricity generating sub-technologies,

K: the set of electricity generating master technologies.

We associate each sub-technology in I with its respective master technology in K using a mapping function: $f : I \rightarrow K$, s.t.

f(Nuclear-light water reactor) = Nuclear,

f(Nuclear-small modular reactor) = Nuclear,

etc.

t: the current time period, $t \in T$,

i: the electricity generating sub-technology, $i \in I$,

k: the electricity generating master technology, $k \in K$.

Decision Variables

N_{it}: the number of electric infrastructure units of technology i that starts to expand in time period t,

CE_{it}: the capacity of generation of technology i that starts to expand in time period t [MW],

CAP_{it}: the installed capacity for electricity generation of technology i in time period t [MW],

EG_{it}: the electricity generation of technology i in time period t [MWh],

EG_t^{Buy}: the electricity purchased to meet demand in time period t [MWh],

CR_{it}: the capacity retirement of technology i in time period t [MW],

W_{it}^{S1}: binary variable. Equals 1 if capacity expansion of technology i in time period t does not exceed the Step 1 upper bound,

W_{it}^{S2}: binary variable. Equals 1 if capacity expansion of technology i in time period t does not exceed the Step 2 upper bound.

Parameters

CAP_{i0}: the installed capacity for electricity generation of technology i in the beginning time period [MW],

ED_t: the electricity demand in time period t [MWh],

IC_{it}: the overnight capital cost of technology i in time period t [2022$/MW],

RC_{it}: the resource cost for electricity generation of technology i in time period t [2022$/Btu],

MVC_{it}: the variable O&M cost of technology i in time period t [2022$MWh],

MFC_{it}: the fixed O&M cost of technology i in time period t [2022$/MW-yr],

BC_t: the electricity price in time period t [2022$/MWh] ,

$BC_t^{Buy,UB}$: the percentage of electricity purchase upper bound in time period t,

$BC_t^{Buy,Rate}$: the electricity purchase growth rate limit in time period t,

η: the inter-regional transmission loss,

GC_i: the total potential capacity of technology i [MW],

UC_{it}: the cumulative planned capacity retirement of technology i in time period t [MW],

AC_{it}: the cumulative planned capacity addition of technology i in time period t [MW],

RT_i: the usage lifetime of technology i,

LT_i: the construction lead time of technology i,

CF_i: the capacity factor of technology i,

HR_i: the heat rate of technology i [Btu/MWh],

OH_i: the annual operating hours of technology i,

RM: the capacity reserve margin required to maintain the electric reliability,

SZ_i: the size of one unit capacity expansion of technology i [MW],

RPS_t: the planned RPS percentage in time period t,

RPS_t^{Rate}: the limit of RPS growth rate in time period t,

rps_{kt}: the planned percentage of electricity generation of master technology k in time period t,

rgs_{kt}: the planned electricity generation of master technology k in time period t [MWh],

CL_t: the planned percentage of clean electricity generation in time period t,

b_i^{RE}: binary parameter. Equals 1 if technology i is a renewable energy, and 0 otherwise,

b_i^{CL}: binary parameter. Equals 1 if technology i is clean energy, and 0 otherwise,

b_i^{OLD}: binary parameter. Equals 1 if technology i was installed in the beginning of the time periods,

DR: the real discount rate,

UB_{kt}^{S1}: the upper bound of capacity expansion Step 1 of master technology k in time period t [MW],

UB_{kt}^{S2}: the upper bound of capacity expansion Step 2 of master technology k in time period t [MW],

$Ratio_{it}^{S2}$: the ratio of the cost adder of capacity expansion Step 2 of technology i in time period t compared to the cost adder of Step 1,

M: a large positive constant.

3.2.2. Deterministic Formulation

A deterministic Mixed Integer Programming (MIP) for the long-term energy planning model is formulated in this subsection. The objective (3.1) minimizes the total cost, including infrastructure cost (C_{INV}), fixed O&M cost (C_{FOM}), variable O&M cost (C_{VOM}), resource cost for electricity generation (C_{RES}), and electricity purchase cost (C_{BUY}).

$$\min \quad C_{Total} = C_{INV} + C_{FOM} + C_{VOM} + C_{RES} + C_{BUY} \tag{3.1}$$

The infrastructure investment costs (3.2) represents the total annualized overnight capital cost for all technology i and time period t.

$$C_{INV} = \sum_{i \in I} \sum_{t \in T} IC_{it}(W_{it}^{S1} + Ratio_{it}^{S2} W_{it}^{S2}) CE_{it}[\frac{DR(1+DR)^{RT_i}}{(1+DR)^{RT_i} - 1}] \tag{3.2}$$

The last term of Eqn (3.2) $[DR(1+DR)^{RT_i}/(1+DR)^{RT_i}-1]$ represents the annualized factor that takes the usage lifetime of technologies into account. The term $(W_{it}^{S1}+Ratio_{it}^{S2}W_{it}^{S2})$ indicates whether a cost adder will be imposed due to the scarcity of labor and resources in a competitive market.

The resource costs associated with fuel-fired power plants, for instance, the cost of utilizing natural gas to operate combustion turbine generators, can be formulated as:

$$C_{RES} = \sum_{i\in I}\sum_{t\in T} HR_i RC_{it} EG_{it} \tag{3.3}$$

The variable O&M costs (3.4) depends on the amount of electricity generation. The costs include water consumption and waste treatment.

$$C_{VOM} = \sum_{i\in I}\sum_{t\in T} MVC_{it} EG_{it} \tag{3.4}$$

However, the fixed O&M costs (3.5) depend on the size of the installed capacity of technology and do not change with the electricity generation. The cost includes labor, materials, and administration.

$$C_{FOM} = \sum_{i\in I}\sum_{t\in T} MFC_{it} CAP_{it} \tag{3.5}$$

The costs of electricity purchase to meet the demand can be represented as:

$$C_{BUY} = \sum_{t\in T} BC_t EG_t^{Buy} \tag{3.6}$$

The installed capacity for electricity generation of technology i in time period t depends on its installed capacity in time period $t-1$, the capacity expansion in time period $t-LT_i$,

the cumulative planned capacity retirement, the capacity retirement due to reaching the usage lifetime, and the unplanned capacity retirement. Hence, the installed capacity for $\forall t \in T$ can be formulated as:

$$CAP_{it} = CAP_{i,t-1} + CE_{i,t-LT_i} + AC_{it} - UC_{it} - CE_{i,t-RT_i-LT_i} - CR_{it} \tag{3.7}$$

$$CAP_{i,t-1} = CAP_{i,t-2} + CE_{i,t-1-LT_i} + AC_{i,t-1} - UC_{i,t-1} - CE_{i,t-1-RT_i-LT_i} - CR_{i,t-1}$$

$$\dots$$

$$CAP_{i1} = CAP_{i0} + CE_{i,1-LT_i} + AC_{i1} - UC_{i1} - CE_{i,1-RT_i-LT_i} - CR_{i1}$$

$CE_{i,t-LT_i}$ represents the capacity expansion decision made in time period $t - LT_i$, and the amount of the capacity is not available until time period t, where LT_i is the lead (construction) time of technology i. Similarly, $CE_{i,t-RT_i-LT_i}$ represents the retired capacity of technology i in time period t due to reaching the lifetime RT_i. The equation can be rewritten as a closed-form formulation:

$$CAP_{it} = CAP_{i0} + \sum_{m=1}^{t-LT_i} CE_{im} + AC_{it} - UC_{it} - \sum_{m=1}^{t-RT_i-LT_i} CE_{im} - CR_{it} \quad \forall i \in I, t \in T \tag{3.8}$$

The constraint (3.9) forces the unplanned capacity retirement to only incur in the capacity already installed at the beginning of the time period. In other words, we avoid retiring the capacity expansion until they reach the usage lifetime.

$$\sum_{t \in T} CR_{it} \leq CAP_{i0} \quad \forall i \in I, t \in T \tag{3.9}$$

The Eq. (3.10) ensures that the capacity expansion is the integer multiple of one unit of installation size.

$$CE_{it} = N_{it} SZ_{it} \quad \forall i \in I, t \in T \tag{3.10}$$

The generation constraint (3.11) represents that the electricity generation of technology i in time period t is limited by the multiplication of operating hours, capacity factor, transmission efficiency $(1-\eta)$, and installed capacity.

$$EG_{it} \leq CF_i OH_i (1-\eta) CAP_{it} \quad \forall i \in I, t \in T \tag{3.11}$$

The reserve margin constraint ensures that the available generating capacity can meet the demand for additional electricity generation, thereby ensuring the reliability of the electricity supply.

$$(1+RM) \sum_{i \in I} EG_{it} \leq \sum_{i \in I} CF_i OH_i (1-\eta) CAP_{it} \quad \forall t \in T \tag{3.12}$$

The constraint (3.13) requires that conventional technologies provide the reserve margin of the capacity, as conventional technologies can generate electricity more consistently than renewable technologies.

$$\sum_{i \in I} (1 - b_i^{RE}) CAP_{it} \geq RM \sum_{i \in I} CAP_{it} \quad \forall t \in T \tag{3.13}$$

The potential capacity constraint (3.14) limits renewable technologies' maximum permissible installed capacity. This limitation emerges because the renewable energy potential capacity is intricately linked to the geographical attributes of the location, including factors such as sunlight duration, wind velocities, river patterns, the presence of a coastline, and so on. Therefore, renewable energy exhibits a natural ceiling on its installed capacity, unlike conventional technologies, due to these geographical attributes.

$$b_i^{RE} CAP_{it} \leq GC_{it} \quad \forall i \in I, t \in T \tag{3.14}$$

The constraint (3.15) ensures that the combined electricity generation and purchase guarantee the demand fulfillment in each time period.

$$\sum_{i \in I} EG_{it} + EG_t^{Buy} \geq ED_t \quad \forall t \in T \tag{3.15}$$

The RPS constraint (3.16) guarantees that renewable technologies produce the planned percentage of electricity generation. The constraint (3.17) imposes the RPS growth rate limitation.

$$\sum_{i \in I} b_i^{RE} EG_{it} \geq RPS_t \sum_{i \in I} EG_{it} \quad \forall t \in T \tag{3.16}$$

$$\sum_{i \in I} b_i^{RE} EG_{it} \leq (1 + RPS_t^{Rate}) \sum_{i \in I} b_i^{RE} EG_{i,t-1} \quad \forall t \in T : t > 1 \tag{3.17}$$

The constraint (3.18) ensures that a specific renewable master technology produces electricity to achieve the targeted electricity generation percentage.

$$\sum_{i \in I : f(i) = k} b_i^{RE} EG_{it} \geq rps_{kt} \sum_{i \in I} EG_{it} \quad \forall k \in K, t \in T \tag{3.18}$$

The constraint (3.19) represents that a targeted amount of electricity generation is produced by a specific renewable master technology.

$$\sum_{i \in I : f(i) = k} b_i^{RE} EG_{it} \geq rgs_{kt} \quad \forall k \in K, t \in T \tag{3.19}$$

The constraint (3.20) describes that a planned percentage of electricity generation is produced by clean energy. The definition of clean energy can differ among states. Typically,

clean energy is defined as technologies that have zero emissions.

$$\sum_{i \in I} b_i^{CL} EG_{it} \geq CL_t \sum_{i \in I} EG_{it} \quad \forall t \in T \tag{3.20}$$

The constraint (3.21) restricts the capacity expansion to technologies that were installed before the time periods commenced. This is attributed to the lack of detailed data on installed technologies. For instance, Annual Energy Outlook 2023 [60] only offers an overview of the installed capacity for the combined cycle without providing a detailed technology classification. A combined cycle can be segmented into categories like single-shaft, multi-shaft, 90% carbon capture and sequestration (CCS), etc. These variations have different impacts on their associated costs, and using average values as replacements is not suitable. Therefore, we limit the capacity expansion to the installed technologies.

$$b_i^{OLD} CE_{it} = 0 \quad \forall i \in I, t \in T \tag{3.21}$$

The constraint (3.22) imposes a maximum percentage for the electricity purchase. Meanwhile, The constraints (3.23) and (3.24) limit the growth rate of the electricity purchase to prevent abrupt declines or surges.

$$EG_t^{Buy} \leq EG_t^{Buy,UB} (\sum_{i \in I} EG_{it} + EG_t^{Buy}) \quad \forall t \in T \tag{3.22}$$

$$EG_t^{Buy} \leq (1 + EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \quad \forall t \in T : t > 1 \tag{3.23}$$

$$EG_t^{Buy} \geq (1 - EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \quad \forall t \in T : t > 1 \tag{3.24}$$

The constraints (3.25) - (3.29) are designed to determine which cost adder will be applied to the capacity expansion for the master technologies. The cost adder represents the additional cost resulting from the competition for labor and resources. Specifically, these

constraints enforce the binary variables $W_{it}^{S1} = 1$ and $W_{it}^{S2} = 0$ if the capacity remains within the threshold of Step 1. Conversely, if the capacity expansion surpasses the threshold of Step 1, these constraints ensure that the binary variables are set to $W_{it}^{S1} = 0$ and $W_{it}^{S2} = 1$.

$$\sum_{i \in I: f(i)=k} CE_{it} \leq UB_{kt}^{S1} + M(1 - W_{it}^{S1}) \quad \forall i \in I : f(i) = k, k \in K, t \in T \tag{3.25}$$

$$\sum_{i \in I: f(i)=k} CE_{it} \geq UB_{kt}^{S1} - M(1 - W_{it}^{S2}) \quad \forall i \in I : f(i) = k, k \in K, t \in T \tag{3.26}$$

$$\sum_{i \in I: f(i)=k} CE_{it} \leq UB_{kt}^{S2} \quad \forall k \in K, t \in T \tag{3.27}$$

$$W_{it}^{S1} + W_{it}^{S2} = 1 \quad \forall i \in I, t \in T \tag{3.28}$$

$$W_{it}^{S1}, W_{it}^{S2} \in \{0, 1\} \quad \forall i \in I, t \in T \tag{3.29}$$

Therefore, the full expression of the deterministic long-term energy planning model **RE** is formulated as follows:

$$\begin{aligned}
\min \quad & C_{INV} + C_{RES} + C_{VOM} + C_{FOM} + C_{BUY} && \textbf{(RE)} \\
\text{s.t.} \quad & C_{INV} = \sum_{i \in I} \sum_{t \in T} IC_{it}(W_{it}^{S1} + Ratio_{it}^{S2} W_{it}^{S2}) CE_{it} [\frac{DR(1+DR)^{RT_i}}{(1+DR)^{RT_i} - 1}] \\
& C_{RES} = \sum_{i \in I} \sum_{t \in T} HR_i RC_{it} EG_{it} \\
& C_{VOM} = \sum_{i \in I} \sum_{t \in T} MVC_{it} EG_{it} \\
& C_{FOM} = \sum_{i \in I} \sum_{t \in T} MFC_{it} CAP_{it} \\
& C_{BUY} = \sum_{t \in T} BC_t EG_t^{Buy} \\
& CAP_{it} = [CAP_{i0} + \sum_{m=1}^{t-LT_i} CE_{im} + AC_{it}
\end{aligned}$$

$$- UC_{it} - \sum_{m=1}^{t-RT_i-LT_i} CE_{im} - CR_{it}] \qquad \forall i \in I, t \in T$$

$$\sum_{t \in T} CR_{it} \leq CAP_{i0} \qquad \forall i \in I, t \in T$$

$$CE_{it} = N_{it} SZ_{it} \qquad \forall i \in I, t \in T$$

$$EG_{it} \leq CF_i OH_i (1 - \eta) CAP_{it} \qquad \forall i \in I, t \in T$$

$$(1 + RM) \sum_{i \in I} EG_{it} \leq \sum_{i \in I} CF_i OH_i (1 - \eta) CAP_{it} \qquad \forall t \in T$$

$$\sum_{i \in I} (1 - b_i^{RE}) CAP_{it} \geq RM \sum_{i \in I} CAP_{it} \qquad \forall t \in T$$

$$b_i^{RE} CAP_{it} \leq GC_{it} \qquad \forall i \in I, t \in T$$

$$\sum_{i \in I} EG_{it} + EG_t^{Buy} \geq ED_t \qquad \forall t \in T$$

$$\sum_{i \in I} b_i^{RE} EG_{it} \geq RPS_t \sum_{i \in I} EG_{it} \qquad \forall t \in T$$

$$\sum_{i \in I} b_i^{RE} EG_{it} \leq (1 + RPS_t^{Rate}) \sum_{i \in I} b_i^{RE} EG_{i,t-1} \qquad \forall t \in T : t > 1$$

$$\sum_{i \in I : f(i) = k} b_i^{RE} EG_{it} \geq rps_{kt} \sum_{i \in I} EG_{it} \qquad \forall k \in K, t \in T$$

$$\sum_{i \in I : f(i) = k} b_i^{RE} EG_{it} \geq rgs_{kt} \qquad \forall k \in K, t \in T$$

$$\sum_{i \in I} b_i^{CL} EG_{it} \geq CL_t \sum_{i \in I} EG_{it} \qquad \forall t \in T$$

$$b_i^{OLD} CE_{it} = 0 \qquad \forall i \in I, t \in T$$

$$EG_t^{Buy} \leq EG_t^{Buy,UB} (\sum_{i \in I} EG_{it} + EG_t^{Buy}) \qquad \forall t \in T$$

$$EG_t^{Buy} \leq (1 + EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \qquad \forall t \in T : t > 1$$

$$EG_t^{Buy} \geq (1 - EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \qquad \forall t \in T : t > 1$$

$$\sum_{i \in I : f(i) = k} CE_{it} \leq UB_{kt}^{S1} + M(1 - W_{it}^{S1}) \qquad \forall i \in I : f(i) = k, k \in K, t \in T$$

$$\sum_{i \in I : f(i) = k} CE_{it} \geq UB_{kt}^{S1} - M(1 - W_{it}^{S2}) \qquad \forall i \in I : f(i) = k, k \in K, t \in T$$

$$\sum_{i \in I: f(i)=k} CE_{it} \leq UB_{kt}^{S2} \qquad \forall k \in K, t \in T$$

$$W_{it}^{S1} + W_{it}^{S2} = 1 \qquad \forall i \in I, t \in T$$

$$W_{it}^{S1}, W_{it}^{S2} \in \{0, 1\} \qquad \forall i \in I, t \in T$$

$$EG_t^{Buy} \geq 0 \qquad \forall t \in T$$

$$N_{it}, CE_{it}, CAP_{it}, EG_{it}, CR_{it} \geq 0 \qquad \forall i \in I, t \in T$$

3.3. Robust Approach

3.3.1. Modeling Uncertainty

In our long-term energy planning model, we consider the modeling of uncertainty that is motivated by [19]. We consider uncertain parameters in the objective function, including the resource cost (RC_{it}), the variable O&M cost (MVC_{it}), the fixed O&M cost (MFC_{it}), and the electricity purchase cost (BC_{it}). The uncertain parameter in the constraint is the electricity demand (ED_t).

Let denote the set containing all uncertain parameters in the objective function as:

$$\Psi = \{RC_{it}, MVC_{it}, MFC_{it}, BC_t\} \; \forall i \in I, t \in T$$

We model $\psi_{it} \in \Psi \; \forall i \in I, t \in T$ takes value in $[\bar{\psi}_{it} - \hat{\psi}_{it}, \bar{\psi}_{it} + \hat{\psi}_{it}]$ where $\bar{\psi}_{it}$ is denoted as the nominal value of the parameter, and $\hat{\psi}_{it}$ is denoted as the deviation of the nominal value. We define z_{it}^{ψ} as the scaled deviation of ψ from its nominal value $\bar{\psi}_{it}$, so that

$$z_{it}^{\psi} = (\psi_{it} - \bar{\psi}_{it})/\hat{\psi}_{it} \in [-1, 1] \; \forall i \in I, t \in T$$

.

The parameter Γ^{ψ} is introduced as a measure for the budget of uncertainty, aiming to control the degree of conservatism. The decision maker chooses a value for Γ^{ψ} to represent the acceptable level of the worst-case scenario. This is represented as:

$$\sum_{i\in I}\sum_{t\in T}|z_{it}^{\psi}|\leq \Gamma^{\psi}$$

We denote $\mathcal{U}^{\psi}$ the uncertainty set for the uncertain parameters in the objective function, which can be defined as:

$$\mathcal{U}^{\psi}=\Big\{\psi\Big|\exists z^{\psi},\psi_{it}=\bar{\psi}_{it}+z_{it}^{\psi}\hat{\psi}_{it}\,\forall i\in I\;t\in T,|z_{it}^{\psi}|\leq 1\,\forall i\in I\;t\in T,\sum_{i\in I}\sum_{t\in T}|z_{it}^{\psi}|\leq\Gamma^{\psi}\,\forall\Big\}\,\forall\psi\in\Psi \tag{3.30}$$

To address uncertainty in electricity demand, our model provides decision-makers the flexibility to control the demand uncertainty in each time period. The underlying rationale is that major planned events often drive demand uncertainty in the upcoming years. For example, constructing infrastructure and accommodating tourists for events like the Olympics or World Cups brings variability in electricity demand in the preparatory period. However, demand uncertainty can be reduced after the events end and return the normal activities. Our model allows the decision-makers to toggle demand uncertainty on or off in each time period. The planner can activate the demand uncertainty in the duration of the events. In other periods with stable demand, uncertainty can be deactivated for a more nominal forecast.

Let $\bar{ED}_t$ denote the nominal demand for all $t\in T$, and denote $\hat{ED}_t$ the deviation from its nominal value for all $t\in T$. We define the scaled deviation $z_t^{ED}=(ED_t-\bar{ED}_t)/\hat{ED}_{it}\in[-1,1]\;\forall i\in I,t\in T$. We introduce the parameter Γ_t^{ED} as the measure for the budget of uncertainty. We define the uncertainty set of the uncertainty in electricity demand as:

$$\mathcal{U}^{ED}=\Big\{ED\Big|\exists z^{ED},ED_t=\bar{ED}_t+z_t^{ED}\hat{ED}_t\;\forall t\in T,|z_t^{ED}|\leq 1\;\forall t\in T, \tag{3.31}$$

$$|z_t^{ED}| \leq \Gamma_t^{ED} \ \forall t \in T \Big\}$$

3.3.2. Robust Formulation

Then, with the constructed uncertainty sets $\mathcal{U}^{\psi}$ and $\mathcal{U}^{ED}$, we formulate the robust optimization problem as:

$$\min \quad \max_{\{\psi \in \mathcal{U}^{\psi}, ED \in \mathcal{U}^{ED}\}} C_{INV} + \tilde{C}_{RES} + \tilde{C}_{VOM} + \tilde{C}_{FOM} + \tilde{C}_{BUY} \tag{3.32}$$

$$\text{s.t.} \quad C_{INV} = \sum_{i \in I} \sum_{t \in T} IC_{it}(W_{it}^{S1} + Ratio_{it}^{S2} W_{it}^{S2}) CE_{it} [\frac{DR(1+DR)^{RT_i}}{(1+DR)^{RT_i} - 1}]$$

$$\tilde{C}_{RES} = \sum_{i \in I} \sum_{t \in T} HR_i \tilde{RC}_{it} EG_{it}$$

$$\tilde{C}_{VOM} = \sum_{i \in I} \sum_{t \in T} M\tilde{V}C_{it} EG_{it}$$

$$\tilde{C}_{FOM} = \sum_{i \in I} \sum_{t \in T} M\tilde{F}C_{it} CAP_{it}$$

$$\tilde{C}_{BUY} = \sum_{t \in T} \tilde{BC}_t EG_t^{Buy}$$

$$CAP_{it} = [CAP_{i0} + \sum_{m=1}^{t-LT_i} CE_{im} + AC_{it} - UC_{it} - \sum_{m=1}^{t-RT_i-LT_i} CE_{im} - CR_{it}] \qquad \forall i \in I, t \in T$$

$$\sum_{t \in T} CR_{it} \leq CAP_{i0} \qquad \forall i \in I, t \in T$$

$$CE_{it} = N_{it} SZ_{it} \qquad \forall i \in I, t \in T$$

$$EG_{it} \leq CF_i OH_i (1 - \eta) CAP_{it} \qquad \forall i \in I, t \in T$$

$$(1 + RM) \sum_{i \in I} EG_{it} \leq \sum_{i \in I} CF_i OH_i (1 - \eta) CAP_{it} \qquad \forall t \in T$$

$$\sum_{i \in I} (1 - b_i^{RE}) CAP_{it} \geq RM \sum_{i \in I} CAP_{it} \qquad \forall t \in T$$

$$b_i^{RE} CAP_{it} \leq GC_{it} \qquad \forall i \in I, t \in T$$

$$\sum_{i \in I} EG_{it} + EG_t^{Buy} \geq \tilde{ED}_t \qquad \forall t \in T$$

$$\sum_{i \in I} b_i^{RE} EG_{it} \geq RPS_t \sum_{i \in I} EG_{it} \qquad \forall t \in T$$

$$\sum_{i \in I} b_i^{RE} EG_{it} \leq (1 + RPS_t^{Rate}) \sum_{i \in I} b_i^{RE} EG_{i,t-1} \qquad \forall t \in T : t > 1$$

$$\sum_{i \in I : f(i) = k} b_i^{RE} EG_{it} \geq rps_{kt} \sum_{i \in I} EG_{it} \qquad \forall k \in K, t \in T$$

$$\sum_{i \in I : f(i) = k} b_i^{RE} EG_{it} \geq rgs_{kt} \qquad \forall k \in K, t \in T$$

$$\sum_{i \in I} b_i^{CL} EG_{it} \geq CL_t \sum_{i \in I} EG_{it} \qquad \forall t \in T$$

$$b_i^{OLD} CE_{it} = 0 \qquad \forall i \in I, t \in T$$

$$EG_t^{Buy} \leq EG_t^{Buy,UB} (\sum_{i \in I} EG_{it} + EG_t^{Buy}) \qquad \forall t \in T$$

$$EG_t^{Buy} \leq (1 + EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \qquad \forall t \in T : t > 1$$

$$EG_t^{Buy} \geq (1 - EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \qquad \forall t \in T : t > 1$$

$$\sum_{i \in I : f(i) = k} CE_{it} \leq UB_{kt}^{S1} + M(1 - W_{it}^{S1}) \qquad \forall i \in I : f(i) = k, k \in K, t \in T$$

$$\sum_{i \in I : f(i) = k} CE_{it} \geq UB_{kt}^{S1} - M(1 - W_{it}^{S2}) \qquad \forall i \in I : f(i) = k, k \in K, t \in T$$

$$\sum_{i \in I : f(i) = k} CE_{it} \leq UB_{kt}^{S2} \qquad \forall k \in K, t \in T$$

$$W_{it}^{S1} + W_{it}^{S2} = 1 \qquad \forall i \in I, t \in T$$

$$W_{it}^{S1}, W_{it}^{S2} \in \{0, 1\} \qquad \forall i \in I, t \in T$$

$$EG_t^{Buy} \geq 0 \qquad \forall t \in T$$

$$N_{it}, CE_{it}, CAP_{it}, EG_{it}, CR_{it} \geq 0 \qquad \forall i \in I, t \in T$$

However, Problem (3.32) is not directly solvable due to the bi-level min max structure. Theorem 3.1 provides steps for achieving a tractable reformulation.

Theorem 3.1 *The deterministic long-term energy planning model with uncertain parameters Problem (3.32) can be reformulated as the following tractable Robust Problem (**RE_RP**).*

$$
\begin{aligned}
\min \quad & \Bigg\{ \sum_{i\in I}\sum_{t\in T} IC_{it}(W_{it}^{S1} + Ratio_{it}^{S2} W_{it}^{S2}) CE_{it}[\frac{DR(1+DR)^{RT_i}}{(1+DR)^{RT_i}-1}] \\
& + \sum_{i\in I}\sum_{t\in T} HR_i \bar{RC}_{it} EG_{it} \\
& + \sum_{i\in I}\sum_{t\in T} M\bar{V}C_{it} EG_{it} \\
& + \sum_{i\in I}\sum_{t\in T} M\bar{F}C_{it} CAP_{it} \\
& + \sum_{t\in T} \bar{BC}_t EG_t^{Buy} \\
& + \Gamma^{RC}\lambda^{RC} + \sum_{i\in I}\sum_{t\in T} V_{it}^{RC} \\
& + \Gamma^{MVC}\lambda^{MVC} + \sum_{i\in I}\sum_{t\in T} V_{it}^{MVC} \\
& + \Gamma^{MFC}\lambda^{MFC} + \sum_{i\in I}\sum_{t\in T} V_{it}^{MFC} \\
& + \Gamma^{BC}\lambda^{BC} + \sum_{t\in T} V_t^{BC} \Bigg\} \qquad \textbf{(RE\_RP)} \\
s.t. \quad & CAP_{it} = [CAP_{i0} + \sum_{m=1}^{t-LT_i} CE_{im} + AC_{it} \\
& \qquad - UC_{it} - \sum_{m=1}^{t-RT_i-LT_i} CE_{im} - CR_{it}] \qquad \forall i \in I, t \in T \\
& \sum_{t\in T} CAP_{it} \le CAP_{i0} \qquad \forall t \in T \\
& CE_{it} = N_{it} SZ_{it} \qquad \forall i \in I, t \in T \\
& EG_{it} \le CF_i OH_i (1-\eta) CAP_{it} \qquad \forall i \in I, t \in T
\end{aligned}
$$

$$(1+RM)\sum_{i\in I} EG_{it} \leq \sum_{i\in I} CF_i OH_i (1-\eta) CAP_{it} \qquad \forall t \in T$$

$$\sum_{i\in I} (1-b_i^{RE}) CAP_{it} \geq RM \sum_{i\in I} CAP_{it} \qquad \forall t \in T$$

$$b_i^{RE} CAP_{it} \leq GC_{it} \qquad \forall i \in I, t \in T$$

$$\sum_{i\in I} EG_{it} + EG_t^{Buy} \geq \bar{ED}_t + \Gamma_t^{ED} \lambda_t^{ED} + \sum_{t\in T} v_t^{ED} \qquad \forall t \in T$$

$$\sum_{i\in I} b_i^{RE} EG_{it} \geq RPS_t \sum_{i\in I} EG_{it} \qquad \forall t \in T$$

$$\sum_{i\in I} b_i^{CL} EG_{it} \geq CL_t \sum_{i\in I} EG_{it} \qquad \forall t \in T$$

$$\sum_{i\in I: f(i)=k} b_i^{RE} EG_{it} \geq rps_{kt} \sum_{i\in I} EG_{it} \qquad \forall k \in K, t \in T$$

$$\sum_{i\in I: f(i)=k} b_i^{RE} EG_{it} \geq rgs_{kt} \qquad \forall k \in K, t \in T$$

$$\sum_{i\in I} b_i^{RE} EG_{it} \leq (1+RPS_t^{Rate}) \sum_{i\in I} b_i^{RE} EG_{i,t-1} \qquad \forall t \in T : t > 1$$

$$b_i^{OLD} CE_{it} = 0 \qquad \forall i \in I, t \in T$$

$$EG_t^{Buy} \leq EG_t^{Buy,UB} (\sum_{i\in I} EG_{it} + EG_t^{Buy}) \qquad \forall t \in T$$

$$EG_t^{Buy} \leq (1+EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \qquad \forall t \in T : t > 1$$

$$EG_t^{Buy} \geq (1-EG_t^{Buy,Rate}) EG_{t-1}^{Buy} \qquad \forall t \in T : t > 1$$

$$\sum_{i\in I: f(i)=k} CE_{it} \leq UB_{kt}^{S1} + M(1-W_{it}^{S1}) \qquad \forall i \in I : f(i) = k, k \in K, t \in T$$

$$\sum_{i\in I: f(i)=k} CE_{it} \geq UB_{kt}^{S1} - M(1-W_{it}^{S2}) \qquad \forall i \in I : f(i) = k, k \in K, t \in T$$

$$\sum_{i\in I: f(i)=k} CE_{it} \leq UB_{kt}^{S2} \qquad \forall k \in K, t \in T$$

$$W_{it}^{S1} + W_{it}^{S2} = 1 \qquad \forall i \in I, t \in T$$

$$\lambda^{RC} + v_{it}^{RC} \geq HR_i \hat{RC}_{it} EG_{it} \qquad \forall i \in I, t \in T$$

$$\lambda^{MVC} + v_{it}^{MVC} \geq M\hat{V}C_{it} EG_{it} \qquad \forall i \in I, t \in T$$

$$\lambda^{MFC} + v_{it}^{MFC} \geq \hat{MFC}_{it} CAP_{it} \qquad \forall i \in I, t \in T$$

$$\lambda^{BC} + v_{t}^{BC} \geq \hat{BC}_{t} EG_{t}^{Buy} \qquad \forall t \in T$$

$$\lambda_{t}^{ED} + v_{t}^{ED} \geq \hat{ED}_{t} \qquad \forall t \in T$$

$$W_{it}^{S1}, W_{it}^{S2} \in \{0,1\} \forall i \in I, t \in T$$

Proof: To linearize the terms $|z| \leq \Gamma$ within the uncertainty sets, introduce

$$z^{+} = \max(0, z) \in [0,1]$$

$$z^{-} = \max(-z, 0) \in [0,1]$$

and it is obvious that

$$z = z^{+} - z^{-}$$

$$|z| = z^{+} + z^{-}$$

Then, the inner-level problem with $\hat{RC}_{it}$ in Problem (3.32) is equivalent to:

$$\begin{aligned} \max_{z^{RC+}, z^{RC-}} \quad & \sum_{i \in I} \sum_{t \in T} (z_{it}^{RC+} - z_{it}^{RC-}) \hat{RC}_{it} HR_i EG_{it} \\ \text{s.t.} \quad & \sum_{i \in I} \sum_{t \in T} (z_{it}^{RC+} + z_{it}^{RC-}) \leq \Gamma^{RC} \\ & z_{it}^{RC+}, z_{it}^{RC-} \leq 1 \quad \forall i \in I, t \in T \\ & z_{it}^{RC+}, z_{it}^{RC-} \geq 0 \quad \forall i \in I, t \in T \end{aligned} \tag{3.33}$$

By applying the method in [19], we obtain the dual formulation of Problem (3.33):

$$\min_{\lambda^{RC}, v^{RC}} \quad \Gamma^{RC} \lambda^{RC} + \sum_{i \in I} \sum_{t \in T} v_{it}^{IC} \tag{3.34}$$

$$\text{s.t.} \quad \lambda^{RC} + v_{it}^{RC} \geq \hat{RC}_{it} HR_i EG_{it} \quad \forall i \in I, t \in T$$
$$\lambda^{RC}, v_{it}^{RC} \geq 0 \quad \forall i \in I, t \in T$$

By utilizing the strong duality in [19], the dual problem (3.34) is feasible and bounded since Problem (3.33) is feasible and bounded. By applying a similar technique, we can derive the dual formulations of other inner-level problems, and substitute the formulations to Problem (3.32). Therefore, we achieve the resulting reformulation provided in Theorem 3.1. □

3.4. Numerical Experiments

In this section, we analyze the results of numerical experiments by using the proposed formulation **RE_RP**. We consider California, consisting of the Western Electricity Coordinating Council - Northern California (CANO) and Western Electricity Coordinating Council - Southern California (CASO). Figure 3.1. shows the components of the electricity market managed by American Electric Reliability Corporation (NERC) or Independent System Operator (ISO) [61]. These regions generally do not align with state boundaries, which complicates the data gathering process since a singular comprehensive data source for the model does not exist. Various resources categorize the data according to different criteria [60,62–64].

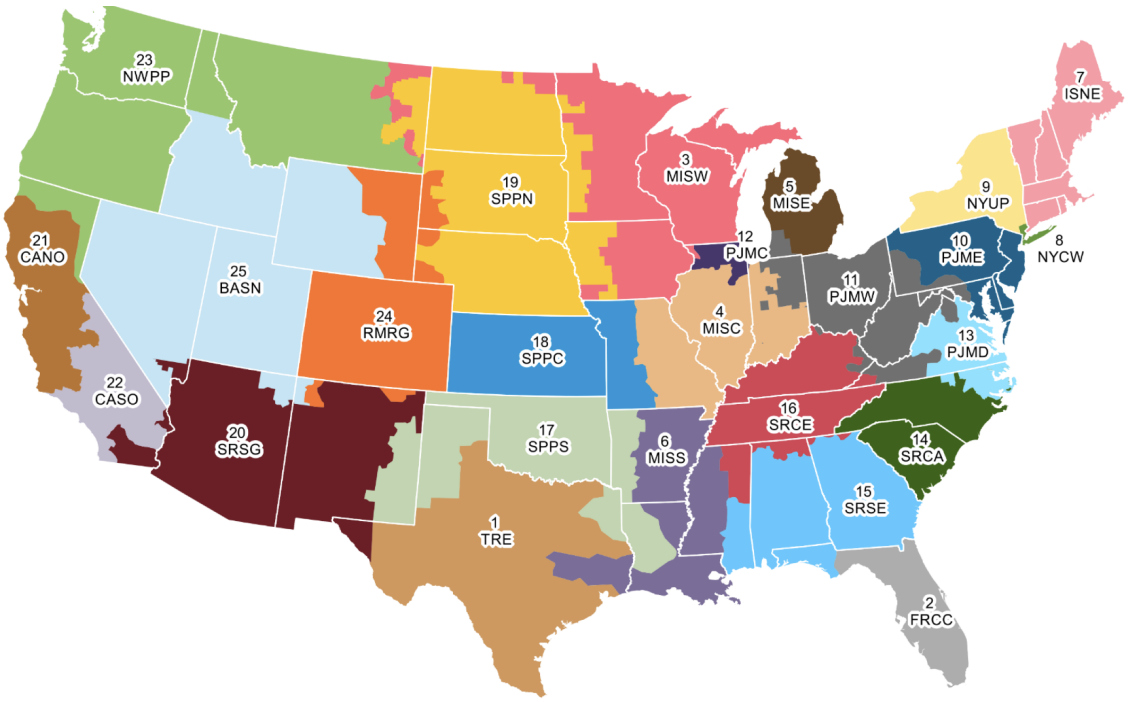

Figure 3.1: U.S. Electricity Market Module Regions
Source: U.S. Energy Information Administration - Electrcity Market Module Regions
https://www.eia.gov/outlooks/aeo/pdf/nerc_map.pdf

Table 3.4: California Environmental Goals

Year	**RPS**	**Clean Electricity**
2024	44%	-
2027	52%	-
2030	60%	90%
2040	-	95%
2045	-	100%

California is chosen as the focus of our numerical experiments for several reasons. The electricity market module regions of CANO and CASO cover a relatively broad geographic

territory within California compared to other states. This extensive coverage ensures that our findings are reflective of a significant portion of the state's infrastructure. In addition, regions of CANO and CASO cover 97.91% operable and standby installed capacity for electricity generation across the entire state. California is one of the states with long-term detailed environmental goals (see Table. 3.4.), and it is consistently updating the public about its progress toward achieving these goals. [65–68].

3.4.1. Setting

We study 141 sub-technologies that can be categorized into 14 master technologies (see Table 3.5). See the detailed technical description of each technology in [69]. Of these, 16 sub-technologies are labeled as 'Old', indicating that their capacity was installed before our study's time periods. The division arises due to the lack of detailed data about the technologies. While EIA 860 [64] provides more segmented information about generators, it often lacks detailed categorization. Since EIA 860 relies on survey-based data collection, many utilities and entities either omit or do not respond comprehensively to several questions. To better understand the impacts of various technologies in long-term energy planning, we categorize them into classes 'Old' and 'New'.

Table 3.5: Overview of Technologies in California

Sub-Technology	**Master Technology (counts)**	b_i^{RE}	b_i^{CL}	b_i^{OLD}
Old_Coal	Coal (1)	0	0	1
Old_NaturalGas_Steam	Natural Gas Steam Turbine (1)	0	0	1
Old_NaturalGas_CombinedCycle	Natural Gas Combined Cycle (1)	0	0	1
Old_NaturalGas_InternalCombustionEngine	Natural Gas Combustion (2)	0	0	1
Old_NaturalGas_CombustionTurbine		0	0	1
Old_Petroleum_InternalCombustionEngine	Petroleum Combustion (2)	0	0	1
Old_Petroleum_CombustionTurbine		0	0	1

Old_Nuclear	Nuclear (1)	0	1	1
Old_FuelCells	Fuel Cell (1)	0	1	1
Old_Hydroelectric	Hydroelectric (1)	1	1	1
Old_Geothermal	Geothermal (1)	1	1	1
Old_WoodWasteBiomass	Biomass (1)	1	1	1
Old_SolarThermal	Solar Thermal (1)	1	1	1
Old_SolarPV	Solar PV (1)	1	1	1
Old_OnshoreWind	Onshore Wind (1)	1	1	1
Old_OffshoreWind	Offshore Wind (1)	1	1	1
New_CominbedCycle_SingleShaft	Natural Gas Combined Cycle (3)	0	0	0
New_CombinedCycle_MultiShaft		0	0	0
New_CombinedCycle_90CCS [1]		0	0	0
New_InternalCombustionEngine	Natural Gas Combustion (3)	0	0	0
New_CombustionTurbine_Aeroderivative		0	0	0
New_CombustionTurbine_IndustrialFrame		0	0	0
New_FuelCells	Fuel Cell (1)	0	1	0
New_Nuclear_LightWaterReactor	Nuclear (2)	0	1	0
New_Nuclear_SmallModularReactor		0	1	0
New_WoodWasteBiomass	Biomass (1)	1	1	0
New_Geothermal	Geothermal (1)	1	1	0
New_Hydroelectric_NPD	Hydroelectric (1)	1	1	0
New_SolarThermal_R1C1 [2]	Solar Thermal (24)	1	1	0
New_SolarThermal_R1C2		1	1	0
New_SolarThermal_R1C3		1	1	0
Continued...		1	1	0
New_SolarPV_R2C1	Solar PV (35)	1	1	0
New_SolarPV_R2C2		1	1	0
New_SolarPV_R2C3		1	1	0

[1] 90CCS: 90% carbon capture and sequestration

[2] The suffixes 'R' and 'C' denote resource and cost classifications, respectively

New_SolarPV_R2C4		1	1	0
New_SolarPV_R2C5		1	1	0
New_SolarPV_R2C6		1	1	0
New_SolarPV_R2C7		1	1	0
Continued...		1	1	0
New_OnshoreWind_R4C1		1	1	0
New_OnshoreWind_R4C2		1	1	0
New_OnshoreWind_R4C3		1	1	0
New_OnshoreWind_R4C4	Onshore Wind (30)	1	1	0
New_OnshoreWind_R4C5		1	1	0
New_OnshoreWind_R4C6		1	1	0
Continued...		1	1	0
New_OffshoreWind_R5C1		1	1	0
New_OffshoreWind_R5C2		1	1	0
New_OffshoreWind_R5C3		1	1	0
New_OffshoreWind_R5C4	Offshore Wind (24)	1	1	0
New_OffshoreWind_R5C5		1	1	0
New_OffshoreWind_R5C6		1	1	0
Continued...		1	1	0

The suffixes 'R' and 'C' in the sub-technologies New Solar Thermal, New Solar PV, New Onshore Wind, and New Offshore Wind represent resource and cost classifications, respectively. For expansive state-level energy planning, applying identical parameters for all solar and wind technologies is not suitable. This is because electricity production from these technologies is significantly influenced by various geographic factors [70–74]. Hence, both solar wind technologies can be classified into different resource categories based on distinct regional attributes, such as solar radiation levels, daylight duration, humidity, and wind velocity. Meanwhile, cost classifications characterize construction complexities arising from

factors of transportation logistics, road conditions, and proximity to the nearest electrical grid [63].

The pre-installed capacity (CAP_{i0}) for technologies is obtained from AEO 2023 [60] and EIA 860 [64,75]. AEO consolidates all combustion technologies, irrespective of their natural gas and petroleum sources, into a single category, which makes the other parameters of combustion technology less accurate. Therefore, we complement this with data from the 'Generator' table of EIA 860, which provides detailed generator-level survey data. Filters are applied to include utility IDs related to the electricity market module regions, generators' status of operation and standby/backup, and all operating years. Subsequently, the prime mover code is filtered to specify the technology type and energy source code to identify natural gas and petroleum.

The capacity factor (CF_i) is estimated by the data provided by various sources [60,63,76]. We assume that capacity factors of generators that primarily use fossil fuels and uranium adhere to the national average [76] since the generation from these generators is not influenced by geographical factors. In contrast, for renewable technologies, we apply Eq (3.35) to calculate the baseline capacity factor in 2022 [60], which specifies the electricity market module regions instead of using a national mean value.

$$CF_i^{Base} = \frac{EG_{it}}{OH_i CAP_{it}} \quad for\ t = 2022 \tag{3.35}$$

where $OH_i = 365$ days/year $\times$ 24 hours/day $= 8760$ hours/year. The computed CF_i^{Base} appears to be lower than the theoretical capacity factor provided by other sources. This difference arises because of our CF_i^{Base} represents the end-use electricity, which takes transmission loss into account. Therefore, we use CF_i^{Base} as the baseline to estimate the capacity factor of sub-technologies characterized by resource and cost classes sourced from EPA's Power Sector Modeling [63]. We compute the average value across all sub-technologies within each master

technology. Subsequently, we determine the ratio between the master technology's capacity factor and CF_i^{Base}, using it to derive the capacity factor for each individual sub-technology.

For the variable and fixed O&M cost rate (MVC_{it},MFC_{it}), heating rate (HR_{it}), size (SZ_i), and construction time (LT_i), we assume they follow the national average value [77]. Nonetheless, the overnight capacity cost rate (IC_{it}) can vary among different electricity market module regions. Given the available data detailing the variation in cost rates across all regions [77], combined with projections of the national average [60], we can deduce the cost rate for CANO&CASO by applying the relevant ratios. A similar approach is applied to technologies with different resources and cost classes. It is worth emphasizing the difference between the nameplate and the real-world installation sizes for solar PV, onshore wind, and offshore wind. To obtain more accurate data, we use the recent 10-year data from 2013 to 2022 to estimate a more realistic average installation size. By summing the cumulative capacity constructed over these ten years and dividing by the total number of respective plants [64], we can estimate the average installation sizes for these technologies.

The projection of electricity demand (ED_t) is given for each electricity market module region [60]. However, the projection of resource cost rate (RC_{it}) is only available for broader census bureau regions, as outlined under 'Energy Prices by Sector and Source' in AEO 2023. Therefore, we use data from the Pacific region, to which California belongs, to represent the resource cost rate in CANO&CASO. The usage lifetime (RT_i) for various technologies are from multiple sources. Specifically, the lifetime for fossil-fueled technologies and biomass is sourced from [63]. The lifetime data for fuel cells is from [78]. The lifetime of nuclear and geothermal technologies is sourced from [60]. Hydroelectricity, solar thermal, solar PV, onshore wind, and offshore wind lifetime are based on information from [79].

We note that the planned capacity retirement (UC_{it}) for technologies such as solar thermal, solar PV, onshore wind, and offshore wind [60], do not reflect the actual expected retirement since their usage lifetime has not been taken into account. To provide a more

accurate representation, we have adjusted these data by adding the respective lifetime to the year of operation.

The upper bound of capacity expansion Step 1 and Step 2 (UB_{kt}^{S1}, UB_{kt}^{S2}), the ratio of cost adder ($Ratio_{it}^{S2}$), and potential capacity for renewable technologies are provided in [63]. We make the following assumptions based on various sources: The upper bound for electricity purchase and its growth rate limit are based on the peak value in California's historical data from 2009 to 2021 [80]. The inter-regional transmission loss within CANO&CASO is set as 4.5% [81]. We assume the capacity reserve margin of 13.9% [63]. The real discount rate is set 3.76% [63]. We assume the RPS growth rate of 15%.

All experiments are implemented using IBM ILOG CPLEX Optimization Studio 22.1.1.0 on the platform:

Model	Apple Macbook Pro 2021
Processor	M1 Max 10-core CPU 3.2 GHz
RAM	64 GB

3.4.2. Results of Deterministic Problem

We first solve the deterministic problem by setting the uncertainty budget of uncertainties, denoted as $\Gamma's$ to zero in Problem **RE_RP**, which is equivalent to solving Problem **RE**. We analyze the trajectory of energy planning over the span from 2023 to 2050 in CANO&CASO. We particularly focus on several critical results: Capacity Expansion (CE_{it}), Electricity Generation (EG_{it}), Installed Capacity (CAP_{it}), and RPS and Clean Electricity Progress.

The evaluation of capacity expansion segmented by different technologies is illustrated in Figure 3.2. In order to prevent being overly congested in the visualization, we consolidated all sub-technologies under New_SolarPV into a single category labeled New_SolarPV_Total.

Similarly, we also merge all sub-technologies under New_OnshoreWind. It reveals that solar PV is a leading energy source for capacity expansion among all renewable technologies, reaching 68.24% of total capacity (see Table 3.6). This observation is intuitive that California has the dominant advantage in solar PV compared to other renewable technologies in terms of a vast solar potential capacity, combined with shorter construction duration and cost efficiency. On the other hand, the non-renewable energy contributors, particularly combustion turbines and combined cycles, contribute 19.86% of total capacity expansion. This observation underlines the role of natural gas as an important energy source in California's long-term energy planning.

Furthermore, the result reveals a significant uptick in capacity expansion during the period from 2043 to 2046. The increase is mainly driven by solar PV, combustion turbines, and nuclear small modular reactors. The observed trend can be attributed to the aging infrastructure that the installations from the earlier periods of the studied time frame approach the end of their technical lifetime and are scheduled for retirement. This mandated retirement leads to an increase in capacity expansion to offset the resulting loss in capacity.

Table 3.6: Total Capacity Expansion (MW) by Tech in CANO&CASO (Deterministic)

Technology	**Total**	**%**
New_SolarPV_Total	72176	68.24%
New_CombustionTurbine_IndustrialFrame	18651	17.63%
New_CombinedCycle_MultiShaft	5527	5.23%
New_Nuclear_SmallModularReactor	4488	4.24%
New_FuelCells	2422	2.29%
New_OnshoreWind_Total	2224	2.10%
New_Hydroelectric_NPD	172	0.16%
New_Geothermal	105	0.10%

The percentage of electricity generation by technologies over the planning horizon is shown in Figure 3.3. Notably, solar PV is a dominant energy contributor to electricity generation. Specifically, 15.63% of the total generation is from solar PV installed prior to the

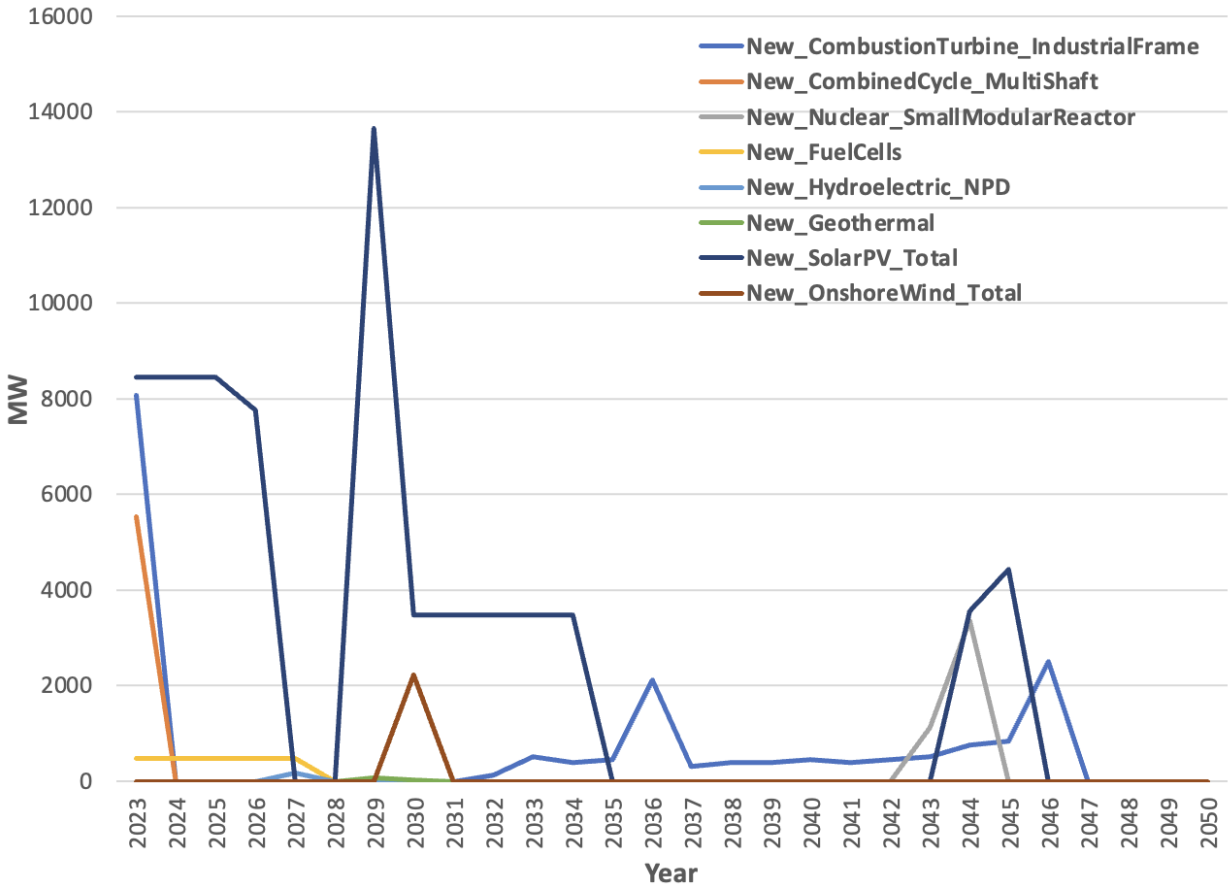

Figure 3.2: Capacity Expansion (MW) by Tech in CANO&CASO (Deterministic)

studied time frame, while 66.21% is from the new solar PV installations. Starting in 2025, we observe a surge in electricity generation from solar PV whose installations were initiated in 2023, factoring in their 2-year construction lead time. As the end of the time frame approaches, there is a decline in generation by solar PV as the early installations reach the end of their usage lifetime. For the pre-installed solar PV, their contribution is relatively consistent. The decline observed in 2050 is also due to reaching the usage lifetime. With the phase-out of new solar PV installations, generation from 'Other Conventional' sources increases to compensate, including 'Old_Coal', 'Old_Nuclear', and 'New_Nuclear_Small Modular Reactor'. Since 'Old' coal-fired and nuclear technologies are forced to retire by 2024 and 2025, respectively, all generation from 'Other Conventional' since 2026 is contributed by 'New_Nuclear_Small Modular Reactor'. However, if we extend the studied time frame, we will not see a sudden drop in solar PV generation in 2050. If the study time frame is extended further, the abrupt drop in solar PV in 2050 would not occur. Instead, we would expect a smoother transition as the new solar PV capacity replaces the retired capacity.

Figure 3.4 shows the installed capacity trajectory of by technology, which aligns a similar trend as electricity generation profiles in Figure 3.3. Natural gas generators are consolidated under the 'NaturalGas' label, with a similar aggregation for other energy sources. The 'Old_NaturalGas' capacity exhibits a downward trend over time. This decline can be attributed to the higher operational costs and planned retirement. In addition, environmental targets further accelerate the phase-out, including 60% RPS and 90% clean electricity goals for 2030. With rising targets for environmental targets after 2030, continued operation of pre-installed fossil fuel plants becomes less economical. The solution retires these facilities over time and expands renewable capacity like solar PV, wind, and nuclear technologies.

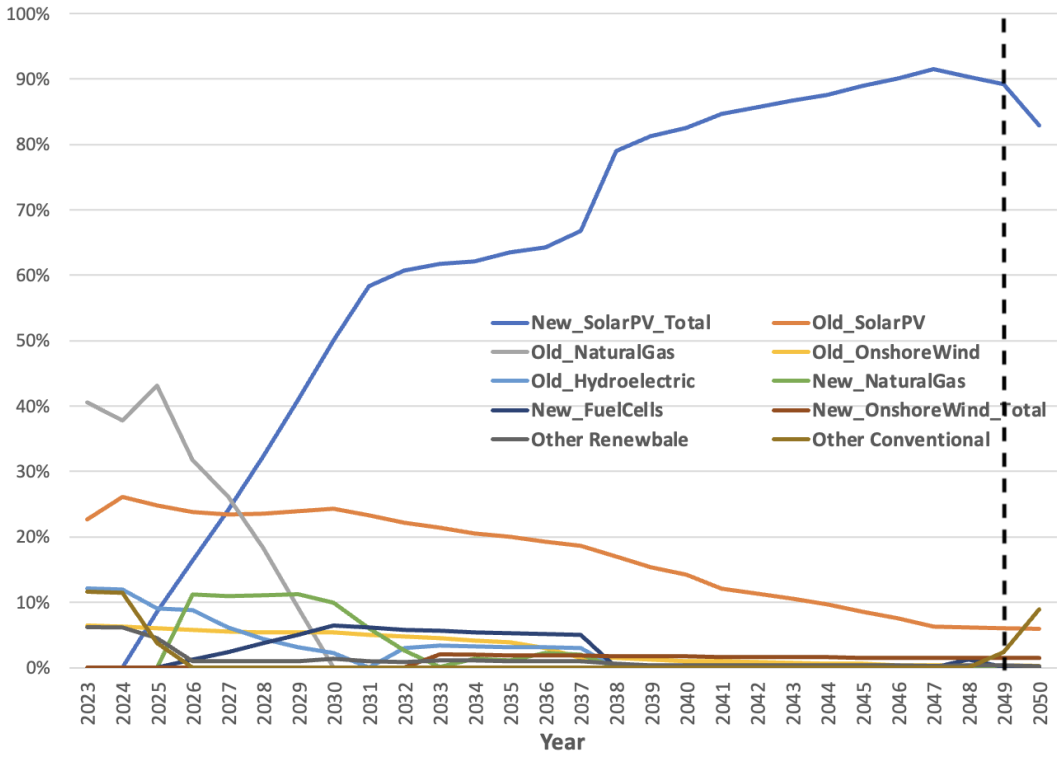

Figure 3.3: Electricity Generation% by Tech in CANO&CASO (Deterministic)

Table 3.7: Total Electricity Generation (MWh) by Tech in CANO&CASO (Deterministic)

Technology	**Total**	**%**
New_SolarPV_Total	5313430477	66.21%
Old_SolarPV	1254439095	15.63%
Old_NaturalGas	465311099	5.80%
Old_OnshoreWind	207421636	2.58%
Old_Hydroelectric	182608051	2.28%
New_NaturalGas	168538075	2.10%
New_FuelCells	152013766	1.89%
New_OnshoreWind_Total	96984413	1.21%
Other Conventional	98801977	1.23%
Other Renewable	85055933	1.06%

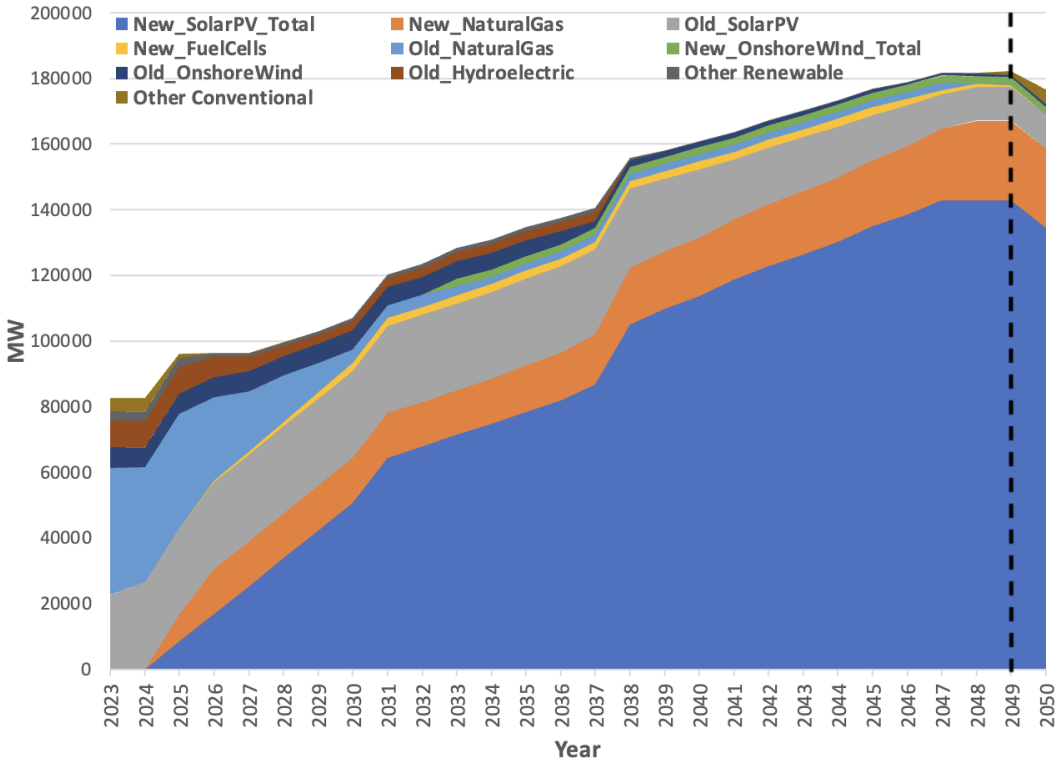

Figure 3.4: Installed Capacity (MW) by Tech in CANO&CASO (Deterministic)

Table 3.8: RPS and Clean Electricity Progress of CANO&CASO (Deterministic)

	2023	2024	2025	2026	2027	2028	2029	2030
RPS%	47.38%	50.32%	52.67%	55.48%	60.00%	66.62%	74.22%	83.24%
Clean Electricity%	55.97%	58.79%	56.86%	57.06%	62.78%	70.68%	79.60%	90.00%
	2031	2032	2033	2034	2035	2036	2037	2038
RPS%	87.78%	91.50%	94.26%	93.15%	93.60%	92.62%	92.90%	100.00%
Clean Electricity%	93.95%	97.34%	99.92%	98.58%	98.90%	97.76%	97.94%	100.00%
	2039	2040	2041	2042	2043	2044	2045	2046
RPS%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Clean Electricity%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	2047	2048	2049	2050				
RPS%	100.00%	98.75%	97.49%	91.08%				
Clean Electricity%	100.00%	100.00%	100.00%	100.00%				

The progress of RPS and clean electricity percentage is shown in Table 3.8. In the initial phase from 2023 to 2030, the RPS percentage steadily increases, starting from 47.38% and surpassing the 60% target to reach above 80% by 2030. The clean electricity percentage achieves the target goal of 90% in 2030 and remains 100% since 2038. The clean electricity percentage achieves the 90% goal by 2030 and maintains 100% from 2038 to 2050.

However, in the later years we notice a decline in the RPS percentage due to the decreased capacity and generation from solar PV. This drop occurs as solar PV, start reaching the end of their usable lifetimes and are retired. Without sufficient replacement by new renewable capacity additions, the RPS percentage decreases as new nuclear technology operates to meet demand while satisfying the clean electricity target. This highlights the importance of continued renewable capacity investment and construction even after short-term RPS milestones are met. Retirements must be anticipated and replaced to maintain the targeted RPS over the long-term planning horizon.

3.4.3. Parameter Uncertainty

In the analysis of the robust optimization problem, we vary values of the budget of the constraint $\Gamma's$ and solved in Problem **RE_RP**, and compare the results from the deterministic problem where $\Gamma's$ are set to zero. The uncertain parameters considered are resource cost rate (RC_{it}), variable O&M cost rate (MVC_{it}), fixed O&M cost rate (MFC_{it}), electricity purchase price (BC_t), and electricity demand (ED_t). These uncertainties are controlled by five $\Gamma's$, respectively:

- $\Gamma^{RC} \in [0, |T| \times \sum_{i \in I} b_i^{RC}]$
- $\Gamma^{MVC} \in [0, |T| \times \sum_{i \in I} b_i^{VC}]$
- $\Gamma^{MFC} \in [0, |T| \times |I|]$

- $\Gamma^{BC} \in [0, |T|]$
- $\Gamma_t^{ED} \in \{0, 1\}\ \forall t \in T$

where b_i^{RC} and b_i^{VC} are additional binary parameters. Set b_i^{RC} to one if the resource cost rate is nonzero and set to zero otherwise. Similarly, set b_i^{VC} to one if the variable O&M cost is nonzero and set to zero otherwise.

Let denote $\Gamma^{p\%}$, the percentage of the highest possible values for the budgets of constraints, and all $\Gamma^{p\%}$'s are rounded to the nearest integers. Additionally, for Γ_t^{ED}, values of one are sequentially assigned in a reverse order. The intuition is that years distant from the current period possess a higher degree of uncertainty than the more recent years. For the scenario $\Gamma^{0\%}$, the problem is equivalent to the deterministic model. For the scenario $\Gamma^{100\%}$, the solution obtained is the most conservative, consistent with Soyster's approach.

In determining the uncertainty parameters, we derive the deviation for the resource cost rate (RC_{it}), electricity purchase price (BC_t), and electricity demand (ED_t) from Annual Energy Outlook 2023 (AEO 2023) [60]. This document provides the reference case for the nominal values, and side cases including 'low economic growth', 'high economic growth', 'low oil price', 'high oil price', 'low oil and gas supply', 'high oil and gas supply', 'low zero-carbon technology cost', 'high zero-carbon technology cost', 'low economic growth and low zero-carbon technology cost', 'high economic growth and low zero-carbon technology cost', 'low economic growth and high zero-carbon technology cost', and 'high economic growth and high zero-carbon technology cost'. See side case descriptions in [82]. On the other hand, for the deviation of variable O&M cost rate (MVC_{it}), and fixed O&M cost rate (MFC_{it}), which are not available in AEO 2023, we determine them from Annual Technology Baseline (ATB2023) [78]. This resource provides three scenarios, namely 'moderate', 'advanced', and 'conservative'.

We first derive the ratio of uncertainty parameters to the nominal values by computing

$$\tilde{\psi}^{Ratio} = \frac{|(\psi^{Max} - \bar{\psi})/\bar{\psi} - (\psi^{Min} - \bar{\psi})/\bar{\psi}|}{2} \tag{3.36}$$

where $\bar{\psi}$ represents the nominal value in AEO 2023 and ATB 2023, termed as reference case or moderate scenario within their data set, respectively. ψ^{Min} and ψ^{Max} represent the lowest and highest values, respectively. Therefore, we estimate uncertain parameters in the model by computing

$$\tilde{\psi} = \tilde{\psi}^{Ratio}\bar{\psi} \tag{3.37}$$

3.4.4. Results of Robust Optimization Problem

In this subsection, we explore how variations in the uncertainty parameter values affect infrastructure investment and energy planning decisions. Figure 3.5 displays the results solved by varying values of Γ's, and their ratio compared to the deterministic scenario ($\Gamma = 0$). A key finding is that the total costs increase gradually from $\Gamma^{0\%}$ to $\Gamma^{70\%}$. However, beyond this threshold, there is a significant spike in costs. This trend is linked to increased uncertainty in the future electricity demand, particularly in more recent years as $\Gamma^{p\%}$'s increase. Consequently, to meet this relatively immediate demand, more infrastructure investments are made, leading to an increase in the O&M costs for these new infrastructures in subsequent years.

Interestingly, the increase in total costs does not correspond to a similar trend in the total capacity expansion as shown in Figure 3.5. This indicates that while total costs increase significantly, the total capacity expansion remains relatively stable. The capacity expansion that occurred earlier in the studied time frame incurs higher fixed O&M costs over time. In addition, as the $\Gamma^{p\%}$ value increases, excessively expanding capacity is not encouraged to

meet the demand due to its associated rising costs. Therefore, the strategy leans towards increasing the utilization of the existing capacity rather than expanding more capacity.

In comparison to the deterministic scenario, the generation from renewable sources gradually increases as $p\%$ surpasses 8%. This trend indicates that when uncertainty in parameters is increased, generation from renewable sources is favored. On the other hand, generation from conventional sources experiences significant fluctuations, reaching its lowest point when $p\%$ is between 30% and 40%, and conventional generation increases again when $p\%$ is beyond this range. This suggests that when uncertainty increases, the utilization of conventional technologies helps to guarantee a reliable energy supply.

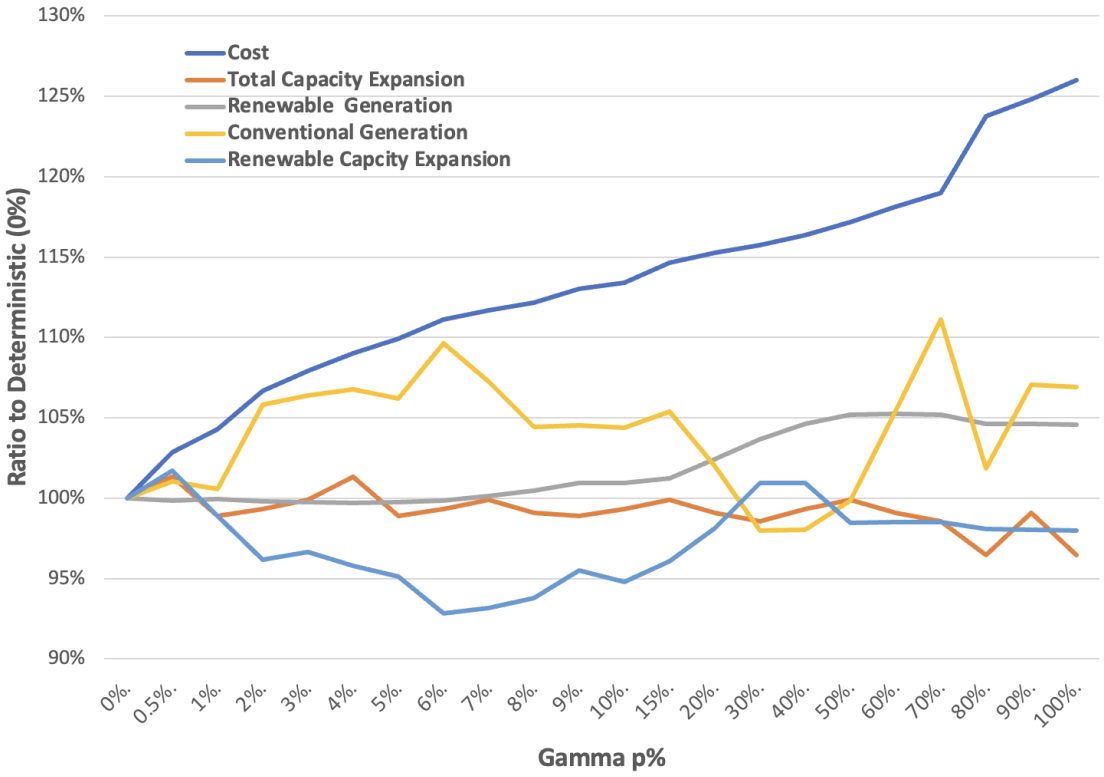

Figure 3.5: Cost, Capacity Expansion, and Generation vs. Deterministic in CANO&CASO

In the study of the total capacity expansion by technology (see Figure 3.6), 33 technologies are consolidated into 10 groups for ease of interpretation. Among these, technologies like the combustion turbine industrial frame and the combined cycle multi-shaft, both affiliated

with natural gas-fueled generators, are labeled as 'New_NaturalGas'. Similarly, all solar PV sub-technologies are categorized into their respective resource class: 3, 4, 5, and 6.

Consistent with findings from the deterministic scenario, solar PV is the leading technology in capacity expansion across all scenarios. However, the variation in capacity allocation among different solar PV resource classes as $\Gamma^{p\%}$ changes. Notably, resource class 6 remains a steady trend across all scenarios. The attractiveness of resource class 6 is underlined by its high capacity factor of 27.36% and relatively low investment costs. From both perspectives of electricity generation efficiency and investment cost, the expansion in solar PV resource class 6 is valuable for all scenarios. However, the actual capacity expansion of resource class 6 is not massive because of its potential capacity, accounting for only 0.84% of the total. Fuel cells and hydroelectric technologies also exhibit a stable trend, with their capacity expansion remaining constant across different scenarios. This consistency indicates their role in the energy mix of meeting clean electricity targets. Interestingly, solar PV resource classes 5 and 6 complement each other, with closely matched capacity factor and potential capacity.

The combined capacity expansion of 'New_NaturalGas' from both combustion turbines and combined cycle technologies shows relatively stable. This stability is crucial in the context of maintaining a capacity reserve margin, which ensures a certain level of installed capacity from conventional technologies for electrical reliability. Additionally, other than renewable technologies, there is an expansion in capacity from fuel cells and nuclear sources. This expansion is part of the strategy to meet the clean electricity target.

The analysis of total electricity generation by technology, as presented in Figure 3.7, shows a similar trend observed in capacity expansion. The leading source of electricity generation across all scenarios is solar PV. In addition, the trend indicates that generation from pre-installed technologies like 'Old_SolarPV', 'Old_OnshoreWind', and 'Old_OtherConventional' remains constant regardless of the scenario. This consistency suggests that pre-installed solar

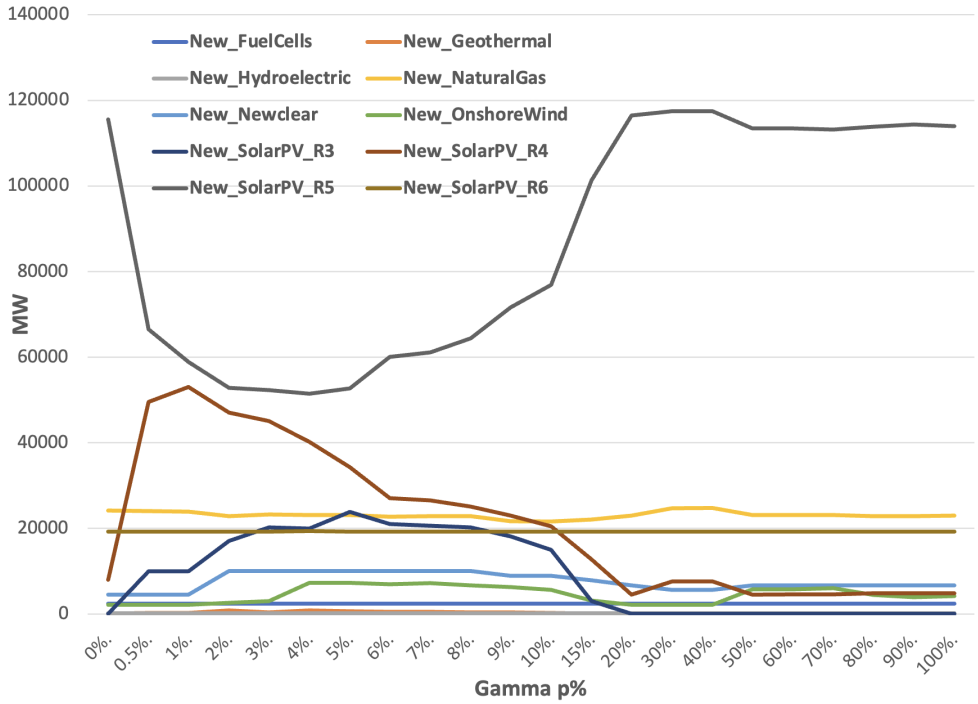

Figure 3.6: Total Capacity Expansion (MW) vs. Deterministic by Tech/Source in CANO&CASO

PV and onshore wind facilities are fully utilized for generation until they reach the technical lifetime.

However, the reason for the stable output from pre-installed conventional technologies is different. Specifically, California's energy policy includes plans to retire all pre-installed coal-powered generators by 2025, nearly 85% of pre-installed steam turbines by 2025, and all pre-installed nuclear capacity by 2026. Additionally, the ambitious goals of RPS and the clean electricity target further reduce the generation from conventional technologies. As a result, there is a noticeable increase in generation from other pre-installed renewable sources as uncertainty increases. This rise is mainly driven by the enhanced utilization of hydroelectric technology.

Figure 3.8 focuses on the percentage of electricity generation by technology. Solar PV consistently emerges as the primary source across all scenarios. The generation from pre-installed solar PV and onshore wind remains constant indicating the full utilization until

their retirement. For the pre-installed conventional plants, similarly observed in Figure 3.7, the low and consistent generation aligns with California's energy policy to retire pre-installed coal-fired and nuclear plants, alongside its planned strict environmental targets.

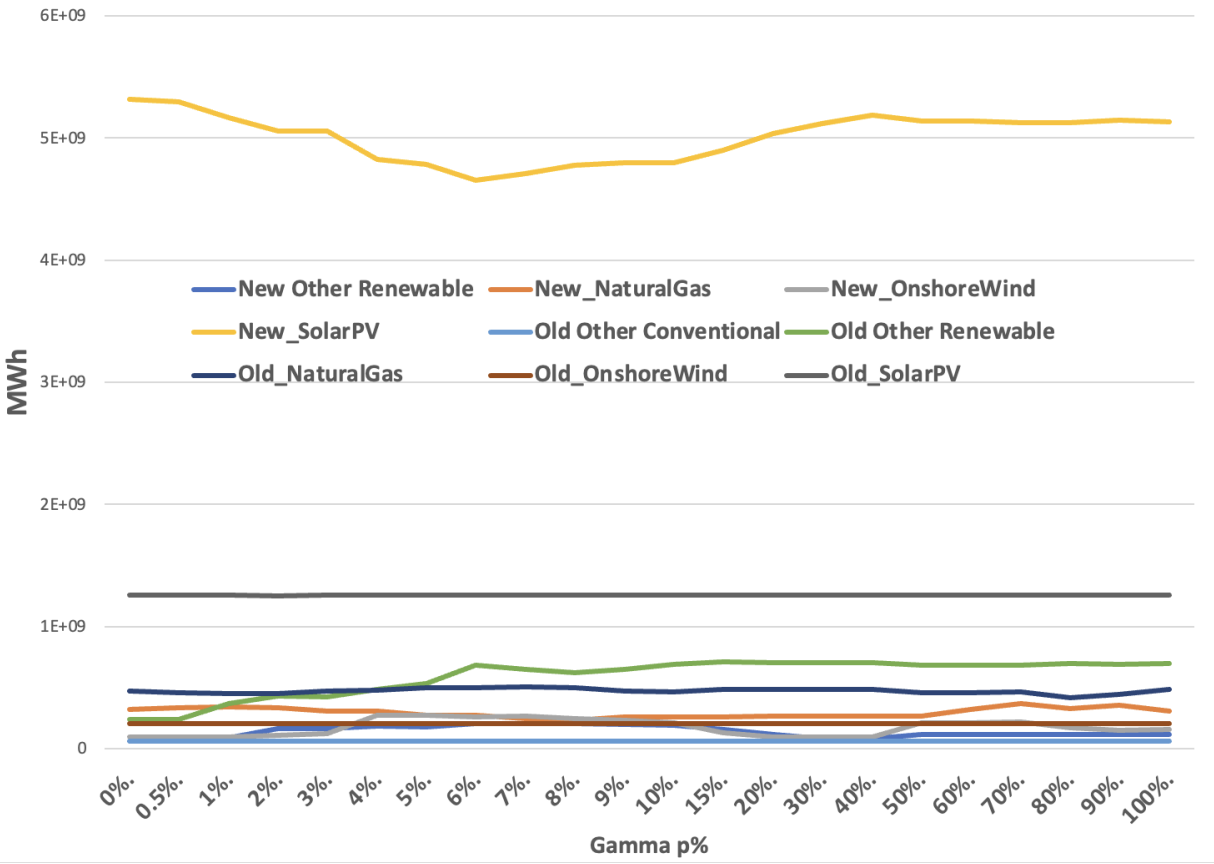

Figure 3.7: Total Electricity Generation (MWh) vs. Deterministic by Tech/Source in CANO&CASO

When analyzing the ratio of generation compared to the deterministic scenario, different technologies show distinct patterns. As $\Gamma^{p\%}$ increases, 'New_Other Renewable' category combined with 'New_Geothermal' and 'New_Hydroelectric' has a notable rise in generation. A similar trend is observed for 'New_OnshoreWind'. Conversely, conventional energy sources like 'New_NatrualGas' exhibit slight fluctuations but generally maintain value close to the deterministic scenario. The 'Old Other Renewable' category, dominated by pre-installed hydroelectric technology, also shows significant increases in generation with higher $\Gamma^{p\%}$ values, which is consistent as observed in Figure 3.7.

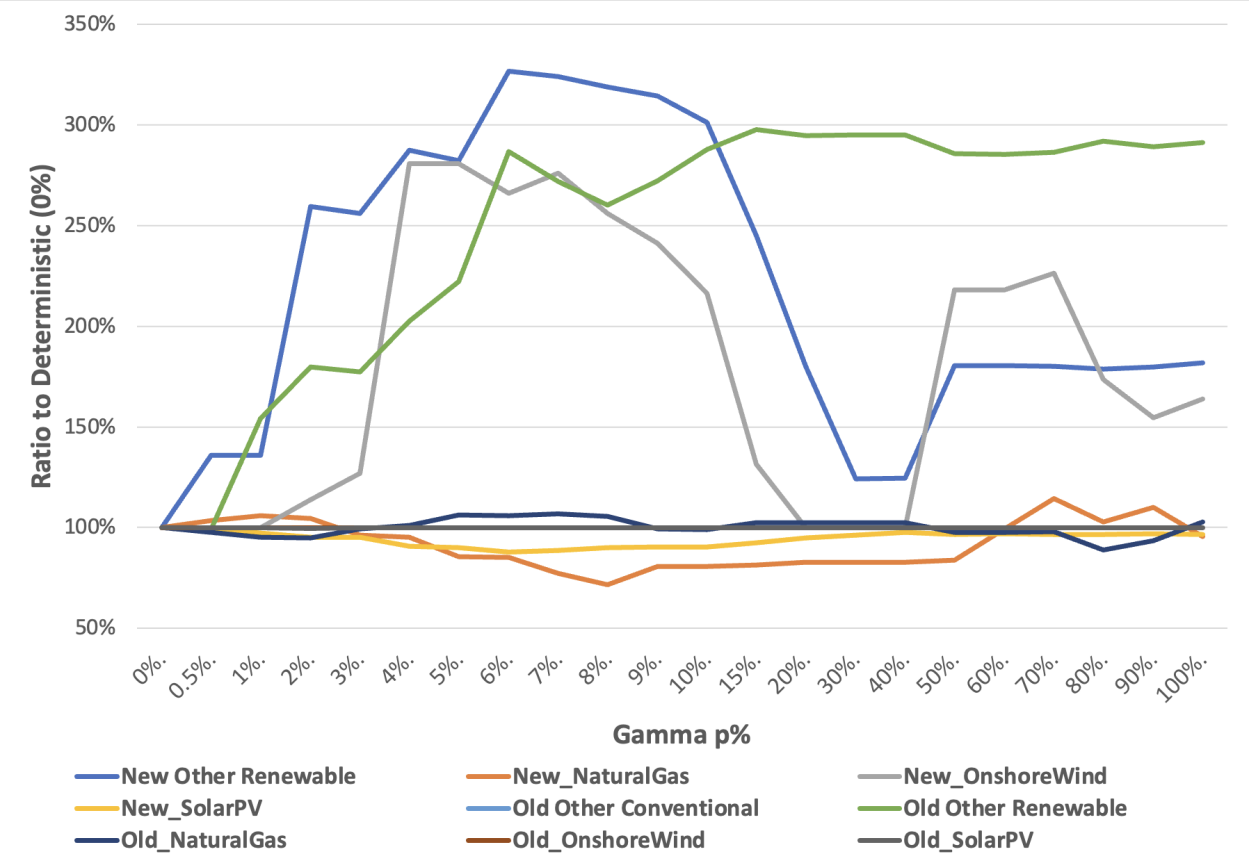

Figure 3.8: Ratio of Total Electricity Generation (MWh) vs. Deterministic by Tech/Source in CANO&CASO

In summary, experiments show that total costs gradually increase from $\Gamma^{0\%}$ to $\Gamma^{70\%}$ followed by a sudden surge due to the growing demand uncertainty. The highest cost is 126% of the deterministic problem at $\Gamma^{100\%}$. The capacity expansion remains stable under higher uncertainty, discouraging capacity expansion at the early phase of the time frame to prevent high operational expenses. Instead, the utilization of pre-installed capacity increases to meet demand. Compared to the deterministic problem, total renewable generation rises as uncertainty increases. Across all scenarios, solar PV is the dominant technology for both capacity expansion and electricity generation. However, the allocation among different solar PV resource classes varies with the values of uncertainty. This variation in allocation reflects the balance of factors such as cost efficiency, capacity factor, and potential capacity.

3.5. Summary

In this chapter, we introduced a long-term energy planning model that integrates renewable energy. The model aims to minimize the total costs incurred over the upcoming decades. These costs include infrastructure investments, the resource cost for electricity generation, the variable and fixed operation and maintenance (O&M) expenses, and electricity purchase while ensuring electrical reliability, complying with potential capacity limitations for various energy sources, meeting electricity demand, and ensuring the planned goals of Renewable Portfolio Standards (RPS) and clean electricity.

We considered the real-world situation in which uncertainty appears in various parameters. Therefore, we incorporated the uncertainty in our model parameters, including resource cost rate, variable and fixed O&M cost rate, electricity price, and demand in the future. We proposed a robust optimization problem and provided a tractable formulation.

We performed numerical experiments on California. Our study spans from 2023 to 2050, a critical period for California to achieve its ambitious goals: 60% RPS by 2030 and 100% RPS by 2045. We vary values of the budget of constraint Γ's to reflect the decision-maker's perspective and provide them with diverse strategies for capacity expansion investment and electricity generation portfolios in the long run.

In the future, we aim to integrate Renewable Energy Certificates (RECs) into our model. RECs serve as exchangeable proofs, certifying the generation and contribution of renewable energy delivered into the power grid [83, 84]. Their significance in renewable energy markets is growing, as many individuals and companies express a willingness to acquire RECs to support renewable energy. This trend could potentially incentivize decision-makers to invest further in renewable energy infrastructure [85]. While numerous studies have delved into the impact of RECs on the market [86–90], there remains a limited amount of literature on the subject. In addition to incorporating RECs, we aim to expand our model to encompass all

states within the nation. Our objective will shift from focusing solely on decision-making for a single state to minimizing the overall social cost for the entire country.

Additionally, we intend to examine the hidden costs associated with renewable energy, which might appear in the coming decades. We may encounter situations where numerous infrastructures approach the end of their usage lifetime, and the recycling of these facilities might incur substantial expenses. Moreover, the potential adverse impacts of renewable installations on the ecology, climate, and various species require further investigation.

Chapter 4

Conclusion

In this dissertation, we investigated the advantages of robust optimization approaches for addressing portfolio selection problem and long-term energy planning model integrated with renewable energy.

In the portfolio allocation problem, we proposed a robust optimization approach for the Mean-Variance portfolio selection problem under uncertainty. The classical Markowitz's Mean-Variable optimization model optimizes portfolio allocation based on expected returns and risks. However, the expected returns are estimated from historical data and are influenced by estimation errors. In this study, we formulate a robust portfolio model using Sharpe's single-index model to describe the stock returns. The Sharpe's model coefficients α and β measure the expected return of a stock, which is independent of market performance, and how sensitive a stock is to the market moves, respectively. We applied Ordinary Least Squares Estimation to compute the coefficients using the historical data. However, uncertainty exists in the coefficients in the real world.

Therefore, we developed a robust optimization model by assuming Sharpe's α and β estimates lie within a prespecified ε from the actual coefficients. The uncertainty set for α and β is motivated by the idea that the actual α and β driving future stock returns to be close to the estimated values. The inherent inaccuracy of estimates calculated from the historical data is potentially due to unmodeled factors. The robust optimization then optimizes the portfolio allocation to minimize the risk while ensuring the minimum expected return from the portfolio based on decision-makers' risk aversion levels. We then provided a tractable reformulation to solve the problem in a commercial solver.

We conducted numerical experiments varying the minimum required expected portfolio return and ε, and analyzed the performance of our model compared to the benchmark. The decision-maker selects ε to reflect his/her attitude towards risk. Experiments of Dow Jones and NASDAQ 100 indices reveal the existence of an optimal solution of $\varepsilon > 0$, achieving a higher return with a lower risk. We also compared our approach with Goldfarb and Iyengar's model. The numerical results indicate that our approach provides the decision-makers with various options for portfolio allocations to match their risk preference. Future research could extend the robust optimization to multi-factor models. Challenges include the construction of an uncertainty set for multiple return factors and tractable reformulation of the associated robust optimization problem.

In the study of renewable energy, we developed a comprehensive long-term energy planning model integrating renewable technologies to meet environmental targets. Unlike typical single-period models assuming immediate capacity availability, our model incorporates realistic construction lead times before new infrastructure becomes operational. We derived installed capacity expression over the planning horizon that captured these construction times. Our model also considered the limited resources in the real world by imposing competition costs for scarce materials and labor in the market.

Furthermore, we formulated a tractable robust optimization considering uncertainty in resource cost, operation and maintenance cost, electricity price, and demand. In addition, the model provided flexibility to control uncertainty in each time period based on decision-makers' plans for future events. In the numerical experiments, we applied the approach to California. The results demonstrated that solar photovoltaic technology is the dominant source for achieving renewable energy goals in California across all scenarios; however, the allocation between solar photovoltaic sub-technologies changed under different scenarios in response to uncertainty. The model optimized solar capacity expansion and generation between its resource and cost classes under different levels of uncertainty. In addition, the study

generated strategies for capacity expansion and electricity generation profiles in California from 2023 to 2050 to achieve its Renewable Portfolio Standards and clean electricity targets.

Future research could incorporate renewable energy credits and incentives and expand the model to the national level. In addition, the costs of recycling the retired renewable infrastructure and the environmental impacts caused by equipment like solar panels are important considerations to include in future studies.

Appendix A

A.1. Chapter 2

Table 1.1: Portfolios Allocation for NDX 86 ($w = 0$)

i \ ε	**0 (Benchmark)**	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
AAPL	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ADBE	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ADI	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ADP	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
ALXN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
AMAT	0	0	0	0	0	0	0	0	0	0	0.0128
AMGN	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
AMZN	0	0	0	0	0	0	0	0	0	0.0149	0.0128
ANSS	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ASML	0	0	0	0	0	0	0	0	0	0.0149	0.0128
ATVI	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
BIIB	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
BMRN	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
CDNS	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CERN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CHKP	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
CMCSA	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
COST	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
CPRT	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
CSCO	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CSGP	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
CSX	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CTAS	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
CTSH	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
CTXS	0	0	0	0	0	0	0	0	0	0	0.0128
DLTR	0	0	0	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
DXCM	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
EA	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
EBAY	0	0	0	0	0	0	0	0	0	0.0149	0.0128
EXC	1	0.5	0.3333	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
EXPE	0	0	0	0	0	0	0	0	0	0	0.0128
FAST	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
FISV	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
GILD	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
GOOGL	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
IDXX	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
ILMN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
INTC	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
INTU	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
ISRG	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
KLAC	0	0	0	0	0	0	0	0	0	0	0.0128
LBTYA	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
LBTYK	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
LRCX	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
LULU	0	0	0	0	0	0	0	0	0	0.0149	0.0128
MAR	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
MCHP	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
MDLZ	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
MNST	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
MSFT	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
MU	0	0	0	0	0	0	0	0	0	0	0.0128
MXIM	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
NFLX	0	0	0	0	0	0	0	0	0	0	0.0128
NTAP	0	0	0	0	0	0	0	0	0	0	0.0128
NTES	0	0	0	0	0	0	0.0345	0.0185	0.0161	0.0149	0.0128
NVDA	0	0	0	0	0	0	0	0	0	0	0.0128
ORLY	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
PAYX	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
PCAR	0	0	0	0	0	0	0	0	0	0.0149	0.0128
PEP	0	0	0.3333	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
QCOM	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
REGN	0	0	0	0	0	0	0	0	0	0	0.0128
ROST	0	0	0	0	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
SBUX	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
SIRI	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
SNPS	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
SWKS	0	0	0	0	0	0	0	0	0.0161	0.0149	0.0128
TCOM	0	0	0	0	0	0	0	0	0	0	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0001	0.0006	0.002	0.005	0.01	0.013	0.03	0.05	0.06	0.1
TMUS	0	0	0	0	0	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
TTWO	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
TXN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
ULTA	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
VRSN	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
VRTX	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
WBA	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128
WDC	0	0	0	0	0	0	0	0	0	0	0.0128
XEL	0	0.5	0.3333	0.2500	0.0714	0.0435	0.0345	0.0185	0.0161	0.0149	0.0128
XLNX	0	0	0	0	0	0	0	0.0185	0.0161	0.0149	0.0128

Table 1.2: Portfolios Allocation for NDX 86 ($w = 0.01$)

i \ ε	**0 (Benchmark)**	0.0005	0.002	0.005	0.011	0.017	0.02	0.04	0.06	0.1
AAPL	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
ADBE	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
ADI	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
ADP	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
ALXN	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
AMAT	0	0	0	0	0	0	0	0	0	0.0128
AMGN	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
AMZN	0	0	0	0	0	0	0	0	0.0149	0.0128
ANSS	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
ASML	0	0	0	0	0	0	0	0	0.0149	0.0128
ATVI	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
BIIB	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
BMRN	0	0	0	0	0	0	0	0	0.0149	0.0128
CDNS	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CERN	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
CHKP	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
CMCSA	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
COST	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
CPRT	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
CSCO	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CSGP	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CSX	0	0	0	0	0	0	0	0.0179	0.0149	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0005	0.002	0.005	0.011	0.017	0.02	0.04	0.06	0.1
CTAS	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
CTSH	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CTXS	0	0	0	0	0	0	0	0	0	0.0128
DLTR	0	0.0853	0.2500	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
DXCM	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
EA	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
EBAY	0	0	0	0	0	0	0	0	0.0149	0.0128
EXC	0.4362	0.4573	0.2500	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
EXPE	0	0	0	0	0	0	0	0	0	0.0128
FAST	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
FISV	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
GILD	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
GOOGL	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
IDXX	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
ILMN	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
INTC	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
INTU	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
ISRG	0	0	0	0	0	0	0	0	0.0149	0.0128
KLAC	0	0	0	0	0	0	0	0	0	0.0128
LBTYA	0	0	0	0	0	0	0	0	0.0149	0.0128
LBTYK	0	0	0	0	0	0	0	0	0.0149	0.0128
LRCX	0	0	0	0	0	0	0	0	0.0149	0.0128
LULU	0	0	0	0	0	0	0	0	0.0149	0.0128
MAR	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
MCHP	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
MDLZ	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
MNST	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
MSFT	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
MU	0	0	0	0	0	0	0	0	0	0.0128
MXIM	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
NFLX	0	0	0	0	0	0	0	0	0	0.0128
NTAP	0	0	0	0	0	0	0	0	0	0.0128
NTES	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
NVDA	0	0	0	0	0	0	0	0	0	0.0128
ORLY	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
PAYX	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
PCAR	0	0	0	0	0	0	0	0	0.0149	0.0128
PEP	0	0	0.2500	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
QCOM	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
REGN	0	0	0	0	0	0	0	0	0	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0005	0.002	0.005	0.011	0.017	0.02	0.04	0.06	0.1
ROST	0	0	0	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
SBUX	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
SIRI	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
SNPS	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
SWKS	0	0	0	0	0	0	0	0	0.0149	0.0128
TCOM	0	0	0	0	0	0	0	0	0	0.0128
TMUS	0	0	0	0	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
TTWO	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
TXN	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
ULTA	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128
VRSN	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
VRTX	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
WBA	0	0	0	0	0	0.0303	0.0244	0.0179	0.0149	0.0128
WDC	0	0	0	0	0	0	0	0	0	0.0128
XEL	0.5638	0.4573	0.2500	0.0714	0.0435	0.0303	0.0244	0.0179	0.0149	0.0128
XLNX	0	0	0	0	0	0	0.0244	0.0179	0.0149	0.0128

Table 1.3: Portfolios Allocation $x_i's$ for NDX 86 ($w = 0.015$)

i \ ε	**0 (Benchmark)**	0.0003	0.0007	0.001	0.002	0.006	0.017	0.04	0.06	0.1
AAPL	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
ADBE	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
ADI	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
ADP	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
ALXN	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
AMAT	0	0	0	0	0	0	0	0	0	0.0128
AMGN	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
AMZN	0	0	0	0	0	0	0	0	0.0149	0.0128
ANSS	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
ASML	0	0	0	0	0	0	0	0	0.0149	0.0128
ATVI	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
BIIB	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
BMRN	0	0	0	0	0	0	0	0	0.0149	0.0128
CDNS	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CERN	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
CHKP	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0003	0.0007	0.001	0.002	0.006	0.017	0.04	0.06	0.1
CMCSA	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
COST	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
CPRT	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
CSCO	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CSGP	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CSX	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CTAS	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
CTSH	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
CTXS	0	0	0	0	0	0	0	0	0	0.0128
DLTR	0	0.4718	0.2957	0.2169	0.1662	0.0588	0.0303	0.0179	0.0149	0.0128
DXCM	0.0655	0	0.1128	0.1323	0.0031	0	0.0303	0.0179	0.0149	0.0128
EA	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
EBAY	0	0	0	0	0	0	0	0	0.0149	0.0128
EXC	0	0.0563	0.2957	0.2169	0.1662	0.0588	0.0303	0.0179	0.0149	0.0128
EXPE	0	0	0	0	0	0	0	0	0	0.0128
FAST	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
FISV	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
GILD	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
GOOGL	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
IDXX	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
ILMN	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
INTC	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
INTU	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
ISRG	0	0	0	0	0	0	0	0	0.0149	0.0128
KLAC	0	0	0	0	0	0	0	0	0	0.0128
LBTYA	0	0	0	0	0	0	0	0	0.0149	0.0128
LBTYK	0	0	0	0	0	0	0	0	0.0149	0.0128
LRCX	0	0	0	0	0	0	0	0	0.0149	0.0128
LULU	0	0	0	0	0	0	0	0	0.0149	0.0128
MAR	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
MCHP	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
MDLZ	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
MNST	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
MSFT	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
MU	0	0	0	0	0	0	0	0	0	0.0128
MXIM	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
NFLX	0	0	0	0	0	0	0	0	0	0.0128
NTAP	0	0	0	0	0	0	0	0	0	0.0128
NTES	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
NVDA	0	0	0	0	0	0	0	0	0	0.0128

Continued from previous page

i \ ε	**0 (Benchmark)**	0.0003	0.0007	0.001	0.002	0.006	0.017	0.04	0.06	0.1
ORLY	0	0	0	0	0.1662	0.0588	0.0303	0.0179	0.0149	0.0128
PAYX	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
PCAR	0	0	0	0	0	0	0	0	0.0149	0.0128
PEP	0	0	0	0.2169	0.1662	0.0588	0.0303	0.0179	0.0149	0.0128
QCOM	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
REGN	0	0	0	0	0	0	0	0	0	0.0128
ROST	0	0	0	0	0.1662	0.0588	0.0303	0.0179	0.0149	0.0128
SBUX	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
SIRI	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
SNPS	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
SWKS	0	0	0	0	0	0	0	0	0.0149	0.0128
TCOM	0	0	0	0	0	0	0	0	0	0.0128
TMUS	0	0	0	0	0	0.0588	0.0303	0.0179	0.0149	0.0128
TTWO	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
TXN	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
ULTA	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
VRSN	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
VRTX	0	0	0	0	0	0	0	0.0179	0.0149	0.0128
WBA	0	0	0	0	0	0	0.0303	0.0179	0.0149	0.0128
WDC	0	0	0	0	0	0	0	0	0	0.0128
XEL	0.9345	0.4718	0.2957	0.2169	0.1662	0.0588	0.0303	0.0179	0.0149	0.0128
XLNX	0	0	0	0	0	0	0	0.0179	0.0149	0.0128

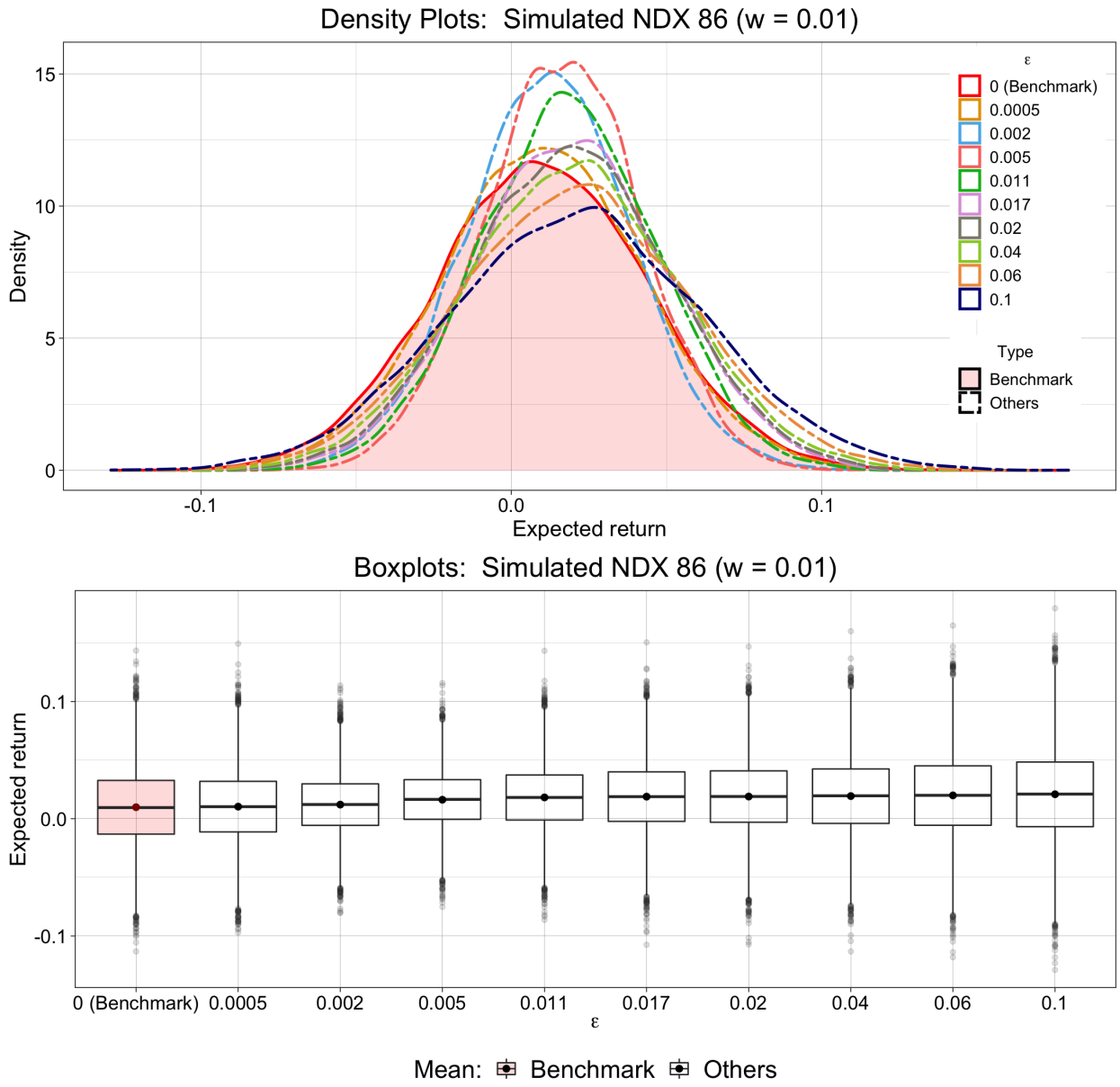

Figure 1.1: Performance on Simulated NDX 86 ($w = 0.01$)

Table 1.4: Simulated NDX 86 ($w = 0.01$)

ε	0	0.0005	0.002	0.005	0.011	0.017	0.02	0.04	0.06	0.1
Mean	0.010	0.010	0.012	0.016	0.018	0.019	0.019	0.019	0.020	0.021
Median	0.009	0.010	0.012	0.016	0.018	0.019	0.019	0.020	0.020	0.021
S.D	0.034	0.032	0.026	0.025	0.028	0.032	0.032	0.034	0.037	0.041

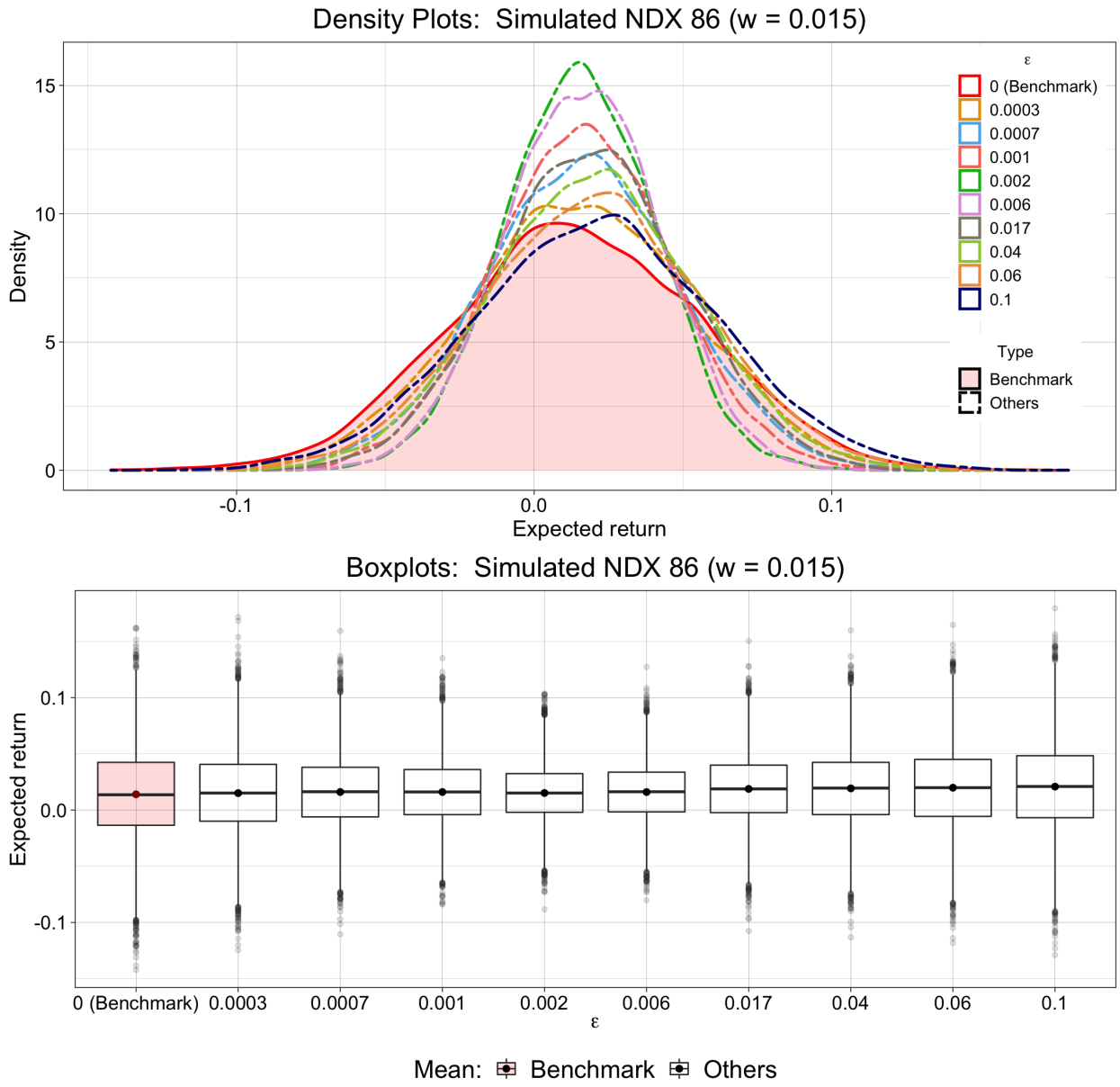

Figure 1.2: Performance on Simulated NDX 86 ($w = 0.015$)

Table 1.5: Simulated NDX 86 ($w = 0.015$)

ε	0	0.0003	0.0007	0.001	0.002	0.006	0.017	0.04	0.06	0.1
Mean	0.014	0.015	0.016	0.016	0.015	0.016	0.019	0.019	0.020	0.021
Median	0.014	0.015	0.016	0.016	0.015	0.016	0.019	0.020	0.020	0.021
S.D	0.041	0.038	0.033	0.030	0.025	0.026	0.032	0.034	0.037	0.041

A.2. Chapter 3

Table 1.6: Electricity Generation (MWh) by Tech in CANO&CASO

$\Gamma^{p\%}$	New Other RE	New_NG	New_OnshoreWind	New_SolarPV	Old Other Conv	Old Other RE	Old_NG	Old_OnshoreWind	Old_SolarPV
0%	63529571	320551841	96984413	5313430477	57762068	239022916	471462505	207421636	1254571405
0.5%	86322187	332108018	96984413	5292899851	57762068	236710337	459815773	207421636	1254571405
1%	86356223	339916175	96984413	5165203019	57762068	368209996	447867073	207421636	1254571405
2%	164907986	334733620	110474116	5058074793	57762068	430024345	447754455	207421636	1254571405
3%	162656808	308879780	123214391	5055642384	57762068	424225665	468649336	207421636	1254571405
4%	182649812	304849605	272428721	4822306106	57762068	483973169	477017556	207421636	1254571405
5%	179270720	274291658	272428721	4784327843	57762068	531446781	501479321	207421636	1254571405
6%	207582659	272959190	257998028	4655952748	57762068	685860161	498451790	207421636	1254571405
7%	205841860	247709988	267618490	4705157133	57762068	649639430	502871136	207421636	1254571405
8%	202611088	228983358	248377565	4778696578	57762068	622141179	498016223	207421636	1254571405
9%	199750631	258150897	233946872	4800375009	57762068	651185320	467780339	207421636	1254571405
10%	191486393	258726393	209895716	4796361708	57762068	688007638	466957961	207421636	1254571405
15%	155653358	260726256	127461149	4902297075	57762068	711912391	483129235	207421636	1254571405
20%	114339549	264769923	96984413	5033290842	57762068	704509691	483093614	207421636	1254571405
30%	78887984	264816993	96984413	5118373308	57762068	705232584	483106248	207421636	1254571405
40%	79034900	264785600	96984413	5186910040	57762068	705073475	483126317	207421636	1254571405
50%	114632662	268829204	211397076	5137646101	57762068	682998660	459793066	207421636	1254571405
60%	114619330	318721499	211397076	5139240941	57762068	682091660	459802014	207421636	1254571405
70%	114468348	366710177	219390974	5128020359	57762068	684966711	462235671	207421636	1254571405
80%	113511998	329300596	168238313	5124242644	57762068	697483871	418410991	207421636	1254571405
90%	114172216	352645151	149996625	5148060602	57762068	691196036	440567057	207421636	1254571405
100%	115502658	306135000	159061727	5132816372	57762068	695924683	484668986	207421636	1254571405

Table 1.7: Total Capacity Expansion (MW) by Tech in CANO&CASO

$\Gamma^{p\%}$	New_FC	New_Geothermal	New_Hydro	New_NG	New_Nuclear	New_OnshoreWind	New_SolarPV_R3	New_SolarPV_R4	New_SolarPV_R5	New_SolarPV_R6
0%.	2422	105	172	24178	4488	2224	0	7999	115594	19342
0.5%.	2422	230	172	24067	4488	2224	9927	49556	66443	19342
1%.	2422	230	172	23868	4488	2224	9927	53090	58822	19342
2%.	2422	784	172	22927	10098	2616	17103	47047	52834	19342
3%.	2422	450	172	23196	10098	3008	20246	45051	52321	19342
4%.	2422	784	172	23099	10098	7325	19976	40235	51485	19356
5%.	2422	659	172	23111	10098	7325	23834	34301	52739	19342
6%.	2422	481	172	22726	10098	6933	20988	27098	60009	19342
7%.	2422	460	172	22885	10098	7194	20691	26491	61129	19342
8%.	2422	429	172	22905	10098	6671	20219	25088	64460	19342
9%.	2422	376	172	21724	8976	6279	18155	22944	71636	19342
10%.	2422	314	172	21596	8976	5625	14985	20543	76856	19342
15%.	2422	63	172	22064	7854	3139	2967	12692	101351	19342
20%.	2422	0	172	23045	6732	2224	0	4559	116417	19342
30%.	2422	0	172	24674	5610	2224	0	7661	117388	19342
40%.	2422	0	172	24738	5610	2224	0	7661	117388	19342
50%.	2422	0	172	23173	6732	5755	0	4559	113409	19342
60%.	2422	0	172	23173	6732	5755	0	4573	113409	19342
70%.	2422	0	172	23109	6732	6017	0	4573	113193	19342
80%.	2422	0	172	22905	6732	4447	0	4842	113854	19342
90%.	2422	0	172	22917	6732	3924	0	4842	114286	19342
100%.	2422	0	172	22961	6732	4186	0	4842	114003	19342

BIBLIOGRAPHY

[1] J. R. Birge and F. Louveaux, Introduction to stochastic programming. Springer Science & Business Media, 2011. 2

[2] A. Ben-Tal, L. El Ghaoui and A. Nemirovski, Robust optimization, vol. 28. Princeton university press, 2009. 2

[3] H. Rahimian and S. Mehrotra, Distributionally robust optimization: A review, arXiv preprint arXiv:1908.05659 (2019) . 3

[4] H. M. Markowits, Portfolio selection, Journal of finance **7** (1952) 71–91. 3, 7

[5] V. K. Chopra and W. T. Ziemba, The effect of errors in means, variances, and covariances on optimal portfolio choice, in Handbook of the fundamentals of financial decision making: Part I, pp. 365–373. World Scientific, 2013. 7

[6] J. L. Treynor, Market value, time, and risk, Time, and Risk (August 8, 1961) (1961) . 7

[7] W. F. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk, The journal of finance **19** (1964) 425–442. 7

[8] J. Lintner, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, The Review of Economics and Statistics **47** (1965) 13–37. 7

[9] J. Mossin, Equilibrium in a capital asset market, Econometrica: Journal of the econometric society (1966) 768–783. 7

[10] F. Black, M. C. Jensen, M. Scholes et al., The capital asset pricing model: Some empirical tests, . 7, 8

[11] E. F. Fama and J. D. MacBeth, Risk, return, and equilibrium: Empirical tests, Journal of political economy **81** (1973) 607–636. 7, 8

[12] M. M. Carhart, On persistence in mutual fund performance, The Journal of finance **52** (1997) 57–82. 8

[13] E. F. Fama and K. R. French, A five-factor asset pricing model, Journal of financial economics **116** (2015) 1–22. 8

[14] A. L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, Operations research **21** (1973) 1154–1157. 8

[15] L. El Ghaoui and H. Lebret, Robust solutions to least-squares problems with uncertain data, SIAM Journal on matrix analysis and applications **18** (1997) 1035–1064. 8

[16] A. Ben-Tal and A. Nemirovski, Robust convex optimization, Mathematics of operations research **23** (1998) 769–805. 8

[17] A. Ben-Tal and A. Nemirovski, Robust solutions of uncertain linear programs, Operations research letters **25** (1999) 1–13. 8

[18] D. Goldfarb and G. Iyengar, Robust portfolio selection problems, Mathematics of operations research **28** (2003) 1–38. 8, 10, 45

[19] D. Bertsimas and M. Sim, The price of robustness, Operations research **52** (2004) 35–53. 8, 19, 21, 33, 68, 74, 75

[20] S. Zymler, B. Rustem and D. Kuhn, Robust portfolio optimization with derivative insurance guarantees, European Journal of Operational Research **210** (2011) 410–424. 8

[21] I. Kakouris and B. Rustem, Robust portfolio optimization with copulas, European Journal of Operational Research **235** (2014) 28–37. 8

[22] A. N. Elmachtoub and P. Grigas, Smart "predict, then optimize", Management Science **68** (2022) 9–26. 8, 9

[23] D. Bertsimas, L. Kogan and A. W. Lo, Hedging derivative securities and incomplete markets: An [epsilon]-arbitrage approach, Operations research **49** (2001) 372–397. 9

[24] S. Chen, Robust option pricing: An [epsilon]-arbitrage approach. PhD thesis, Master's thesis, MIT, 2010. 9

[25] C. Bandi and D. Bertsimas, Robust option pricing, European Journal of Operational Research **239** (2014) 842–853. 9

[26] J. H. Lee and L. Jiao, On quasi [epsilon]-solution for robust convex optimization problems, Optimization Letters **11** (2017) 1609–1622. 10

[27] J. J. Buonocore, P. Luckow, G. Norris, J. D. Spengler, B. Biewald, J. Fisher et al., Health and climate benefits of different energy-efficiency and renewable energy choices, Nature Climate Change **6** (2016) 100–105. 51

[28] R. Wiser, D. Millstein, T. Mai, J. Macknick, A. Carpenter, S. Cohen et al., The environmental and public health benefits of achieving high penetrations of solar energy in the united states, Energy **113** (2016) 472–486. 51

[29] O. Ellabban, H. Abu-Rub and F. Blaabjerg, Renewable energy resources: Current status, future prospects and their enabling technology, Renewable and sustainable energy reviews **39** (2014) 748–764. 51

[30] N. Lior, Energy resources and use: The present situation and possible paths to the future, Energy **33** (2008) 842–857. 51

[31] B. Hadjerioua, Y. Wei and S.-C. Kao, An assessment of energy potential at non-powered dams in the united states, tech. rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 2012. 51

[32] A. Brooks, Renewable energy resource assessment information for the united states, tech. rep., EERE Publication and Product Library, Washington, DC (United States), 2022. 51

[33] A. Lopez, B. Roberts, D. Heimiller, N. Blair and G. Porro, Us renewable energy technical potentials. a gis-based analysis, tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012. 51

[34] G. L. Barbose, Us renewables portfolio standards 2021 status update: Early release, tech. rep., Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2021. 51

[35] A. Nogee, J. Deyette and S. Clemmer, The projected impacts of a national renewable portfolio standard, The Electricity Journal **20** (2007) 33–47. 51

[36] R. H. Wiser, G. Barbose, J. Heeter, T. Mai, L. Bird, M. Bolinger et al., A retrospective analysis of the benefits and impacts of us renewable portfolio standards, . 51

[37] M.-Y. Huang, J. R. Alavalapati, D. R. Carter and M. H. Langholtz, Is the choice of renewable portfolio standards random?, Energy Policy **35** (2007) 5571–5575. 52

[38] H. Yin and N. Powers, Do state renewable portfolio standards promote in-state renewable generation?, Energy Policy **38** (2010) 1140–1149. 52

[39] M. Fischlein and T. M. Smith, Revisiting renewable portfolio standard effectiveness: policy design and outcome specification matter, Policy Sciences **46** (2013) 277–310. 52

[40] J. Joshi, Do renewable portfolio standards increase renewable energy capacity? evidence from the united states, Journal of Environmental Management **287** (2021) 112261. 52

[41] T. Mai, R. Wiser, G. Barbose, L. Bird, J. Heeter, D. Keyser et al., A prospective analysis of the costs, benefits, and impacts of us renewable portfolio standards, tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2016. 52

[42] W. Short, P. Sullivan, T. Mai, M. Mowers, C. Uriarte, N. Blair et al., Regional energy deployment system (reeds), tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2011. 52

[43] J. Ho, J. Becker, M. Brown, P. Brown, I. Chernyakhovskiy, S. Cohen et al., Regional energy deployment system (reeds) model documentation: Version 2020, tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2021. 52

[44] J. Ding and A. Somani, A long-term investment planning model for mixed energy infrastructure integrated with renewable energy, in 2010 IEEE green technologies conference, pp. 1–10, IEEE, 2010. 52

[45] C. L. Lara, D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh and I. E. Grossmann, Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm, European Journal of Operational Research **271** (2018) 1037–1054. 52

[46] C. Li, A. J. Conejo, P. Liu, B. P. Omell, J. D. Siirola and I. E. Grossmann, Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems, European Journal of Operational Research **297** (2022) 1071–1082. 52

[47] J. M. Mulvey, R. J. Vanderbei and S. A. Zenios, Robust optimization of large-scale systems, Operations research **43** (1995) 264–281. 52

[48] J. Koo, K. Han and E. S. Yoon, Integration of ccs, emissions trading and volatilities of fuel prices into sustainable energy planning, and its robust optimization, Renewable and Sustainable Energy Reviews **15** (2011) 665–672. 52, 53

[49] K. Akbari, M. M. Nasiri, F. Jolai and S. F. Ghaderi, Optimal investment and unit sizing of distributed energy systems under uncertainty: A robust optimization approach, Energy and Buildings **85** (2014) 275–286. 52, 53

[50] A. T. Rezvan, N. S. Gharneh and G. Gharehpetian, Robust optimization of distributed generation investment in buildings, Energy **48** (2012) 455–463. 53

[51] A. Parisio, C. Del Vecchio and A. Vaccaro, A robust optimization approach to energy hub management, International Journal of Electrical Power & Energy Systems **42** (2012) 98–104. 53

[52] A. Lorca and X. A. Sun, Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind, IEEE Transactions on Power Systems **30** (2014) 1702–1713. 53

[53] M. Zugno and A. J. Conejo, A robust optimization approach to energy and reserve dispatch in electricity markets, European Journal of Operational Research **247** (2015) 659–671. 53

[54] P. Xiong and C. Singh, Distributionally robust optimization for energy and reserve toward fa low-carbon electricity market, Electric Power Systems Research **149** (2017) 137–145. 53

[55] D. Yu, T. Zhang, G. He, S. Nojavan, K. Jermsittiparsert and N. Ghadimi, Energy management of wind-pv-storage-grid based large electricity consumer using robust optimization technique, Journal of Energy Storage **27** (2020) 101054. 53

[56] H. Ranjbar, S. H. Hosseini and H. Zareipour, A robust optimization method for co-planning of transmission systems and merchant distributed energy resources, International Journal of Electrical Power & Energy Systems **118** (2020) 105845. 53

[57] S. Bahramara, Robust optimization of the flexibility-constrained energy management problem for a smart home with rooftop photovoltaic and an energy storage, Journal of Energy Storage **36** (2021) 102358. 53

[58] J. Yang and C. Su, Robust optimization of microgrid based on renewable distributed power generation and load demand uncertainty, Energy **223** (2021) 120043. 53

[59] S. Moret, F. Babonneau, M. Bierlaire and F. Maréchal, Decision support for strategic energy planning: A robust optimization framework, European Journal of Operational Research **280** (2020) 539–554. 53

[60] U.S.EIA, Annual energy outlook 2023 with projections to 2050, tech. rep., U.S. Department of Energy, 2023. 65, 75, 80, 81, 89

[61] U.S.EIA, Electricity market module regions, tech. rep., U.S.Department of Energy, 2023. 75

[62] E. Emissions, Generation resource integrated database (egrid), 2021, Available from EPA's eGRID web site: https://www. epa. gov/egrid (2021) . 75

[63] U.S.EPA, Documentation for epa's power sector modeling platform v6 using the integrated planning model 2022 reference case, tech. rep., U.S.Environmental Protection Agency, 2023. 75, 80, 81, 82

[64] U.S.EPA, Form eia-860 detailed data with previous form data, tech. rep., U.S.Environmental Protection Agency. https://www.eia.gov/electricity/data/eia860/, 2023. 75, 77, 80, 81

[65] NCSL, State renewable portfolio standards and goals, tech. rep., National Conference of State Legislatures https://www.ncsl.org/energy/state-renewable-portfolio-standards-and-goals, 2021. 77

[66] G. L. Barbose, Us state renewables portfolio & clean electricity standards: 2023 status update, tech. rep., Lawrence Berkeley National Laboratory, 2023. 77

[67] California˙State, Senate bill no. 100, chapter 312. an act to amend sections 399.11, 399.15, and 399.30 of, and to add section 454.53 to, the public utilities code, relating to energy, . 77

[68] CES, 2021 sb 100 joint agency report - achieving 100 percent clean electricity in california: An initial assessment, tech. rep., California Energy Commission, 2021. 77

[69] U.S.EIA, Capital cost and performance characteristic estimates for utility scale electric power generating technologies, tech. rep., U.S.Depart of Energy, 2020. 77

[70] T. Yunus Khan, M. E. M. Soudagar, M. Kanchan, A. Afzal, N. R. Banapurmath, N. Akram et al., Optimum location and influence of tilt angle on performance of solar pv panels, Journal of Thermal Analysis and Calorimetry **141** (2020) 511–532. 79

[71] P. Hevia-Koch and J. Ladenburg, Where should wind energy be located? a review of preferences and visualisation approaches for wind turbine locations, Energy Research & Social Science **53** (2019) 23–33. 79

[72] L. C. Rodman and R. K. Meentemeyer, A geographic analysis of wind turbine placement in northern california, Energy Policy **34** (2006) 2137–2149. 79

[73] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin and J. Shelby, The national solar radiation data base (nsrdb), Renewable and sustainable energy reviews **89** (2018) 51–60. 79

[74] C. Draxl, B. Hodge, A. Clifton and J. McCaa, Overview and meteorological validation of the wind integration national dataset toolkit, tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2015. 79

[75] U.S.EPA, Form eia-860 instructions annual electric generator report, tech. rep., U.S.Environmental Protection Agency., 2023. 80

[76] U.S.EIA, Electric power monthly, tech. rep., U.S.Energy Information Administration https://www.eia.gov/electricity/monthly/, 2023. 80

[77] U.S.EIA, Cost and performance characteristics of new generating technologies, annual energy outlook 2023, tech. rep., U.S.Energy Information Administration https://www.eia.gov/outlooks/aeo/assumptions/pdf/elec_cost_perf.pdf, 2023. 81

[78] B. Mirletz, L. Vimmerstedt, T. Stehly, S. Akar, D. Stright, C. Augustine et al., 2023 annual technology baseline (atb) cost and performance data for electricity generation technologies, . 81, 89

[79] IRENA, Renewable power generation costs in 2019, tech. rep., International Rewnable Energy Agency, 2020. 81

[80] CEC, Total system electric generation 2009-2021 with totals ada, tech. rep., California Energy Commission https://www.energy.ca.gov/media/7311, 2022. 82

[81] U.S.EPA, The emissions and generation resource integrated database - egrid technical guide with year 2021 data, tech. rep., Office of Atmospheric Programs Clean Air Markets Division, 2023. 82

[82] U.S.EIA, Annual energy outlook 2023: Case descriptions, tech. rep., U.S. Department of Energy, 2023. 89

[83] U.S.EPA, Renewable energy certificate (rec) arbitrage, tech. rep., U.S.Environmental Protection Agency, 2018. 96

[84] T. Jones, R. Quarrier, M. Kelty et al., The legal basis for renewable energy certificates, Center for Resource Solutions. Retrieved May **25** (2015) 2021. 96

[85] L. Tawney, M. Sotos and E. Holt, Describing purchaser impact in u.s. voluntary renewable energy markets, tech. rep., U.S.Environmental Protection Agency, 2018. 96

[86] E. Holt, J. Sumner and L. Bird, Role of renewable energy certificates in developing new renewable energy projects, tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2011. 96

[87] E. Holt and L. Bird, Emerging markets for renewable energy certificates: opportunities and challenges, . 96

[88] Q. Zhu, X. Chen, M. Song, X. Li and Z. Shen, Impacts of renewable electricity standard and renewable energy certificates on renewable energy investments and carbon emissions, Journal of environmental management **306** (2022) 114495. 96

[89] D. Hulshof, C. Jepma and M. Mulder, Performance of markets for european renewable energy certificates, Energy Policy **128** (2019) 697–710. 96

[90] G. Shrimali and S. Tirumalachetty, Renewable energy certificate markets in india—a review, Renewable and Sustainable Energy Reviews **26** (2013) 702–716. 96

[91] O. M. Rouhani, D. Niemeier, H. O. Gao and G. Bel, Cost-benefit analysis of various california renewable portfolio standard targets: Is a 33% rps optimal?, Renewable and Sustainable Energy Reviews **62** (2016) 1122–1132.

[92] D. Bertsimas and M. Sim, Robust discrete optimization and network flows, Mathematical programming **98** (2003) 49–71.