TUM School of Management
Technische Universität München

Analytics for Healthcare: Critical Care, Emergency and Pandemic Management

Zhiyuan Lou

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigten Dissertation.

Vorsitz:

Prof. Dr. Gudrun Kiesmüller

Prüfende der Dissertation:

1. Prof. Dr. Jingui Xie
2. Prof. Dr. Rainer Kolisch
3. Prof. Dr. Xiaolei Xie

Die Dissertation wurde am 05.05.2025 bei der Technischen Universität München eingereicht und durch die TUM School of Management am 15.09.2025 angenommen.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Jingui Xie, for his invaluable guidance and exceptional technical support throughout my doctoral journey. I am especially grateful for the numerous opportunities he provided for me to engage in international collaborations, which have been pivotal in broadening my academic horizons. I also wish to thank my coauthors, Prof. Dr. Melvyn Sim, Prof. Dr. Chenyi Fu, and Prof. Dr. Zhi Chen. The insightful research discussions and exchanges with them have significantly enriched my knowledge and understanding in the field of operations research.

I am equally grateful to all my colleagues at the Chair of Business Analytics and friends from other teams, who made our office ("L.2.14") and our center a lively and pleasant environment. I am especially thankful to Maotong, whose support during the most difficult times and whose technical discussions were invaluable. My sincere thanks also go to Han and Nan for their assistance with teaching responsibilities, and to Zihan, who always joined our group lunches and generously shared snacks. I would also like to thank other friends, including Dai and Sai, for the enjoyable and refreshing conversations that brightened many of my days.

On a personal level, I want to thank my family for supporting me to pursue my studies (work) abroad. I am also grateful to my new friends in Germany. Special thanks go to Ting-Yu for the delicious cookies and several heartfelt conversations we shared. I cherish the weekend trip with Chris and Nana — Chris's excellent driving and their warm companionship made those moments truly memorable. I am also greatly thankful to Shengkai, whose cooking is the best Chinese food in Baden-Württemberg, at least in our minds. I would also like to thank friends who temporarily stayed in Heilbronn, including Feier, Rong, Ruilu, and Yan, for the wonderful time we spent together and the memorable photos we took. Altogether, I deeply appreciate the unforgettable moments shared with all these friends, especially those from "L.2.14", which made my journey all the more special.

Finally, as John Allen Paulos aptly said, " Uncertainty is the only certainty there is, and knowing how to live with insecurity is the only security." Life may be difficult to optimize, but I believe what we can do is find optimality in our beliefs. I wish all my friends the very best in their future endeavors.

Abstract

This dissertation develops robust and risk-aware decision-making frameworks tailored to healthcare and public health applications, where uncertainty is inherent and often unavoidable. Decision-makers in settings such as intensive care units (ICUs), pandemic response planning, and emergency medical services (EMS) operations routinely face highly uncertain environments characterized by incomplete information, stochastic dynamics, and rapidly evolving conditions. Suboptimal decisions in these critical domains can lead to adverse patient outcomes, resource shortages, and systemic inefficiencies. Traditional optimization models, which typically rely on deterministic or expected-value criteria, often fail to adequately capture these uncertainties. As a result, they may produce solutions that exhibit a high risk of constraint violations and resource misallocations under real-world variability.

Motivated by these challenges, this dissertation systematically integrates robust optimization, robust satisficing, and risk-sensitive optimization principles into three complementary and application-driven models, each addressing uncertainty at different levels of healthcare operations.

The first part of the dissertation introduces a risk-sensitive optimal stopping framework for sequential treatment decisions in ICUs, such as extubation and discharge timing. By embedding the entropic risk measure into a Markov Decision Process (MDP), the model accounts for both expected performance and the variability of adverse outcomes. The resulting threshold-based policies enable an explicit trade-off between patient safety and resource efficiency. A comparative analysis is performed against traditional risk-neutral MDPs, and the framework is further extended to incorporate predicted patient states. Structural properties of the optimal policy under this extended model are also characterized.

The second part focuses on public health-level decision-making by developing a robust optimization framework for pandemic intervention planning under epidemiological uncertainty. The model integrates stochastic compartmental dynamics and introduces a novel ambiguity tolerance metric to quantify the risk of violating key healthcare performance targets, including infection peaks, hospital capacities, and intervention costs. To ensure computational tractability, a bilinear approximation based on asymptotic Gaussian properties is developed. The model's effectiveness is demonstrated through a case study based on real-world data from Singapore during the COVID-19 pandemic.

The third part proposes a unified robust decision-making framework that integrates robust optimization and robust satisficing paradigms, with an emphasis on EMS applications. In this framework, two distinct layers of uncertainty are modeled: outcome uncertainty, describing deviations of realized inputs from specified historical inputs, and estimation uncertainty, which captures discrepancies between specified inputs and the true but unobserved latent variables. A novel Minimum Volume Confidence Set (MVCS) approach is proposed for constructing norm-specific outcome disparity metrics from data. The framework accommodates asymmetry in uncertainty and integrates affine predictive models to incorporate side information into the robust decision-making process. The effectiveness of the framework is demonstrated through a numerical study in the context of EMS resource allocation.

Overall, this dissertation contributes to the development of interpretable, data-driven, and risk-aware decision-support tools for healthcare and public health operations. The proposed methodologies address fundamental challenges related to risk management, uncertainty modeling, and data limitations, and provide actionable insights for both academic research and practical implementation.

Contents

List of Figures

List of Tables

Acronyms

The following table lists the abbreviations used in this thesis.

Abbreviation	Description
CVaR	Conditional Value at Risk. 16, 17, 32, 34, 36
DM	Decision maker. 19, 22, 27, 28
EMS	Emergency medical service. 2, 7, 8, 91, 122, 151
Eq.	Equation. 39, 40, 42–44, 48, 49
FSD	First-order stochastic dominance. 17, 22, 24, 47, 53
GRC	Globalized Robust Counterpart. 11
ICU	Intensive Care Unit. 1, 5
KL	Kullback-Leibler. 11
LHS	Left hand side. 40, 41
LOS	Length of Stay. 3, 7, 13–15, 35, 58
MDP	Markov Decision Process. 3, 5, 6, 8, 9, 14, 15, 19, 24, 30, 36, 37, 63, 151
MICU	Medical Intensive Care Unit. 35, 56, 59
MIMIC-IV	Medical Information Mart for Intensive Care IV. viii, 35, 56, 57, 59
MLR	Monotone likelihood ratio. 24
MVCS	Minimum Volume Confidence Set. 4, 152
NOSP	Nominal optimal stopping problem. 20, 22, 23, 25, 26, 31, 32, 34–36, 47–49
NOSP-P	NOSP with predicted future states. 33, 34, 36, 55
RHS	Right hand side. 40
RLOS	Remaining length of stay. 30, 31
RO	Robust Optimization. 1, 3, 9, 10, 91
ROSP	Risk-sensitive optimal stopping problem. 20–23, 25–29, 31, 32, 34–36, 47–49
ROSP-P	ROSP with predicted future states. 27–29, 33, 34, 36, 51
RS	Robust Satisficing. 10, 11

Abbreviation	Description
SAA	Sample Average Approximation. 9
SOSP	Satisficing optimal stopping problem. 54
SOSP-P	SOSP with predictive information. 55
TP2	Total positivity of order 2. 24, 25, 28, 50–52
VaR	Value at Risk. 16, 17, 32, 34, 36

1 Introduction

The COVID-19 pandemic marked a watershed moment for modern healthcare systems, exposing vulnerabilities in decision-making at every level—from individual patient care to hospital administration and public policy. The speed and severity with which the virus propagated demanded swift, data-informed decisions in the face of immense uncertainty. Governments worldwide implemented non-pharmacological interventions (NPIs)—including lockdowns, quarantines, and social distancing—to mitigate viral spread. However, these policies often came at a steep economic and societal cost, disrupting industries, education, and labor markets globally (ESCAP, et al., 2020).

The extensive impact of the pandemic has underscored the necessity for robust, flexible, and data-driven decision-making frameworks that can function under uncertainty. Healthcare systems, especially in high-stakes settings like Intensive Care Unit (ICU), must navigate a precarious balance: making timely decisions with limited information while safeguarding patient outcomes and resource availability. Decisions such as when to extubate a patient, discharge them from the hospital, or allocate limited ventilators and ICU beds are fraught with uncertainty and can carry life-or-death consequences.

Robust decision-making under uncertainty is not a novel concept (Ben-Tal and Nemirovski, 1998; El Ghaoui, et al., 1998; Bertsimas and Sim, 2004); however, its role in healthcare has gained renewed urgency. Classical approaches to medical decision-making — such as expected utility maximization — often assume perfect knowledge of model parameters and outcome distributions. These assumptions are likely to fail while facing complex real-world medical scenarios. Uncertainty arises from both limited data, rapidly evolving patient or population states and new variants.

This dissertation focuses on the robust decision-making and risk-sensitive frameworks to tackle healthcare decision problems. We utilize Robust Optimization (RO) and risk measure tools to analyze decisions in three kinds of problems: optimal stopping in critical care, pandemic controls and resource allocation in emergency medical service.

Background

Decision-making in healthcare is inherently multidisciplinary, drawing from clinical expertise, statistical modeling, and operational management. Over the past two decades, advancements in health informatics, machine learning, and optimization theory have contributed to the development of prescriptive analytics models that not only predict outcomes but recommend actionable policies. These systems aim to support clinicians and administrators in making better decisions under uncertainty.

First, optimal stopping time in critical care operates at the patient level, where clinicians must determine when to terminate or adjust treatments such as mechanical ventilation or hospitalization. These decisions require balancing the clinical benefits of continued intervention against the risks and costs associated with premature or delayed stopping. For example, extubating a patient too early can result in extubation failure and re-intubation, increasing the risk of mortality (Epstein, et al., 1997; Thille, et al., 2013), while delaying extubation can lead to ventilator-associated pneumonia, airway trauma, and respiratory muscle fatigue. Likewise, premature hospital discharge may increase the likelihood of readmission, whereas delayed discharge leads to inefficient use of critical care resources (Joynt, Jha, et al., 2012; Zuckerman, et al., 2016).

Secondly, developing a prescriptive analytics model to guide government policymakers on the optimal levels of these interventions is challenging. One major difficulty lies in designing intervention policies based on forecasts of the pandemic's near-term trajectory (for instance, over the next few months). Compartmental models have long been studied since the early twenties century (Ross, 1916; Kermack and McKendrick, 1927). Extensions of it have been proposed in recent years, for instance, R. Li, et al. (2020) incorporate the symptomatic and asymptomatic infected individuals; (Chang, et al., 2021) integrate the mobility network and data to study spatial features. However, they are still not able to accurately predict future trend, especially in several months. Regarding this, it is necessary to develop a decision-making framework that captures the uncertainty in disease progression.

Thirdly, the Emergency medical service (EMS) problem centers on optimizing the allocation and deployment of life-saving resources—such as ambulances and ventilators—to ensure timely responses to emergencies, a task made complex by uncertainty in both demand and travel conditions. Traditional models, such as the Location Set Covering Problem (LSCP) and the Maximal Covering Location Problem (MCLP), focus on deterministic coverage of demand points within acceptable response times (Marianov and ReVelle, 1996), but fall short in addressing the stochastic realities of EMS operations. In response, recent research has embraced probabilistic and stochastic modeling approaches that better capture demand variability and travel-time fluctuations, with Beraldi, et al. (2004) introducing a stochastic programming model incorporating probabilistic coverage constraints. These models, however, often rely on the estimation of uncertain distributions, raising concerns about estimation accuracy. Data-driven and machine learning approaches have emerged to improve demand forecasting, particularly in low- and middle-income settings (X. Li, et al., 2011), but estimation uncertainty can degrade decision quality. To mitigate this, RO has been employed to improve solution reliability under uncertainty, as exemplified by Boutilier and T. Chan (2020), who integrate predictive modeling with simulation and real-world data to evaluate EMS deployment policies aimed at reducing response times.

1.1 Contribution and structure

The remainder of this thesis is organized as follows:

Chapter 2 provides a comprehensive review of recent research on healthcare problems and methodologies, including risk optimization in Markov Decision Process (MDP), RO and robust satisficing.

Chapter 3 formulates a risk-sensitive optimal stopping framework by embedding the entropic risk measure into a MDP. This allows the model to account not only for expected values, such as expected Length of Stay (LOS), but also for outcome variability and extreme risks, which are often hidden in average-based evaluations. For instance, while the expected readmission cost might appear low, tail risks—such as long ICU stays for a small percentage of patients—can have significant implications (Shrank, et al., 2019). The proposed framework addresses these concerns and supports more robust clinical policies. We establish that the risk-sensitive optimal policy admits a threshold structure and compare them with those from traditional MDP. There are some comparison results under certain conditions. Furthermore, we incorporate predicted states (information) into our model and analyze the resulting structural properties. Two numerical studies — extubation and discharge — have been done to support our results.

Chapter 4 develops a robust optimization framework for public health decision-making during pandemics, addressing the limitations of deterministic models in capturing the uncertainty of disease spread and healthcare demand. The proposed model leverages a stochastic compartmental structure and introduces a novel *ambiguity tolerance* metric to quantify the risk of violating healthcare performance targets such as peak infections, hospital capacity, and implementation costs. To enhance tractability, the model uses the entropic risk measure and exploits the asymptotic Gaussian properties of compartment populations, enabling a recursive reformulation with bilinear terms. This framework offers a flexible, data-driven tool for policymakers to design resilient and adaptive intervention strategies.

Chapter 5 introduces a unified framework for robust decision-making that integrates both robust optimization and robust satisficing paradigms while addressing two key sources of uncertainty: *outcome uncertainty* (deviation of realized input from the specified input) and *estimation uncertainty* (deviation of the specified input from the true latent value). The construction of two layers of uncertainties plays an important role on the performance of robust decisions in problems like EMS. The chapter proposes a novel Minimum Volume Confidence Set (MVCS) approach to construct norm-specific uncertainty disparity metric from historical data, minimizing conservativeness while maintaining statistical validity. The framework is extended to support asymmetric disparity metrics and allow for the integration of side information through predictive models. For constructing estimation disparity metrics, we adopt a likelihood-based method inspired by T. Zhu, et al. (2022). A numerical study demonstrates the effectiveness of the MVCS approach and unified framework in handling outliers, capturing data asymmetry, and improving performance in resource allocation compared to standard sample-average methods.

Chapter 6 concludes the thesis by summarizing the main findings and insights derived from the three major studies presented.

1.2 Notations

Matrices and vectors are denoted using boldface uppercase and lowercase letters, respectively, such as $\mathbf{P}$ and $\boldsymbol{q}$. The set of real numbers (resp., non-negative real numbers) is denoted by $\mathbb{R}$ (resp., $\mathbb{R}_+$), while the set of integers (resp., non-negative integers) is denoted by $\mathbb{Z}$ (resp., $\mathbb{Z}_+$). We denote by $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$) the space of $n \times n$ positive semidefinite (positive definite) matrices, *i.e.*, $\mathbb{S}^n_+ = \{\mathbf{D} \in \mathbb{R}^{n\times n} | \mathbf{D} \succeq \mathbf{0}\}$. The Frobenius inner product between two matrices is denoted by $\bullet$. Random variables are indicated by a tilde, for example, $\tilde{x} \sim \mathbb{P}$, where $\mathbb{P} \in \mathcal{P}_0$ and $\mathcal{P}_0(\mathcal{Z})$ denotes a specified set of probability distributions on a set $\mathcal{Z}$. Stochastic order relations, such as stochastic dominance and monotone likelihood ratio dominance, are expressed as $\tilde{x} \preceq_{\text{st}} \tilde{y}$ and $\tilde{x} \preceq_{\text{lr}} \tilde{y}$, respectively, for random variables. For vectors representing the probability mass functions of discrete random variables, these relations are written as $\boldsymbol{p} \preceq_{\text{st}} \boldsymbol{q}$ and $\boldsymbol{p} \preceq_{\text{lr}} \boldsymbol{q}$, respectively. The expectation operator is denoted by $\mathbb{E}[\cdot]$. For N a positive integer, $[N] = \{1, 2, 3, \ldots, N\}$ denotes the running indices. The operator $x^+ = \max\{x, 0\}$ (resp., $x^- = \max\{-x, 0\}$) signifies the positive (resp., negative) part of a variable, and can be applied element-wise to vectors, *i.e.*, $\boldsymbol{v}^+$.

2 State of the art

2.1 Healthcare applications

2.1.1 Optimal stopping in critical care

The optimal stopping problem, a fundamental concept with applications in diverse fields such as finance (Boyarchenko and Levendorskii, 2007; Sturt, 2023), healthcare (Alagoz, et al., 2004), and marketing (Feng and Gallego, 1995; Erat and Kavadias, 2008), has been widely studied. Recent research has focused on improving value function approximations and developing efficient, interpretable policies (Goldberg and Y. Chen, 2018; Ciocan and Mišić, 2022). In this section, we narrow our focus to the literature of its application in **critical care settings**, particularly within the ICU, where timely and optimal decision-making is crucial for patient outcomes.

Alagoz, et al. (2004) address the challenge of determining the optimal timing for living-donor liver transplantation to maximize patient outcomes, such as quality-adjusted life expectancy. They develop an MDP model to characterize the evolution of patient health states and derive conditions under which the optimal transplantation policy exhibits a control-limit (or threshold) structure. Similarly, C. W. Chan, et al. (2012) use finite-horizon, discrete-time MDPs to derive optimal ICU policies for discharging patients early and evaluate the performance of an intuitive policy: discharging the patient with the shortest remaining LOS.

In radiation therapy, the concept of Optimal Stopping in Radiation Therapy (OSRT) has been introduced to identify patients for whom continued treatment might be ineffective or even harmful. By modeling patient responses as stochastic processes, OSRT aims to personalize treatment plans, potentially halting therapy when predicted benefits are minimal, thereby sparing patients from unnecessary side effects and allowing for alternative treatments (Ajdari, et al., 2019). In the treatment of rheumatoid arthritis, optimal stopping approaches are used to decide whether to continue or cease medication, to achieve disease remission while minimizing adverse effects (Kotas, 2019). Additionally, Nasrollahzadeh and Khademi (2020) have developed computationally efficient methods to determine optimal stopping points in adaptive Phase II dose-finding clinical trials, where decision-makers might terminate trials early due to demonstrated efficacy or abandon them when futility is evident. El Hassan, et al. (2022) apply discrete-time Markov chain models to ascertain the optimal lockdown duration, balancing the health benefits of reduced transmission against the socio-economic costs of prolonged restrictions. Cheng, et al. (2024) develop a discrete-time, finite-horizon MDP with predictions of future states to support extubation decisions, characterize the structure of the optimal policy, and provide insights into how predictive information can lead to different decision protocols.

Our study, introduced in Chapter 3, is closely related to the work of Cheng, et al. (2024); however, while their research concentrates on the application of predictive information in decision-making, ours emphasizes accounting for the variability of performance outcomes. We also show that our proposed framework can be extended to models with predicted states.

2.1.2 Pandemic control

Pandemic control analysis involves two key components: forecasting disease spread and optimizing intervention strategies. In the following, we review the literature in these areas.

2.1.2.1 Pandemic prediction

Forecasting infection trends is a well-established field, with epidemiological models playing a central role since the early 20th century (Ross, 1916; Kermack and McKendrick, 1927). Various methodologies, including Bayesian estimation (Greenland, 2007), machine learning (Gu, 2020), Markov chains (Lewnard, et al., 2014), time-series regression (Ma, et al., 2022), and agent-based models (Grefenstette, et al., 2013), have been developed to model disease dynamics. Among these, compartmental models like the SEIR (Susceptible, Exposed, Infectious, Removed) framework are widely used. The SEIR model uses ordinary differential equations to predict the average population in each compartment over time.

The COVID-19 pandemic, as the most recent and widespread global health crisis, has driven extensions in compartmental modeling. For example, R. Li, et al. (2020) extend the SEIR model to include symptomatic and asymptomatic cases, using Wuhan data to estimate parameters. Chang, et al. (2021) improve predictions by integrating mobility data, while Bertsimas, et al. (2021a) propose a 13-compartment model for COVID-19. However, these deterministic models often overlook uncertainty, limiting their use in optimization. Stochastic models, which incorporate random parameters like transmission rates (X. Zhang and K. Wang, 2014; Q. Liu, et al., 2017), offer more realism but are analytically complex, hindering their integration into prescriptive frameworks (Lekone and Finkenstädt, 2006).

Despite progress, accurate long-term predictions remain challenging due to factors like viral mutations, population mobility, and environmental changes, complicating decision-making during pandemics.

2.1.2.2 Pandemic decision-making

Building on the predicted numbers from these epidemiological models, a growing body of literature proposes prescriptive analytics models to guide pandemic decision-making. One critical area is the allocation of scarce resources, including tests, vaccines, ventilators, and ICU beds. The distribution procedure must consider factors such as exposure risk and vulnerability to distinct population groups. L. Sun, et al. (2014) develop a static model to allocate patients to hospitals based on resource availability and LOS dependencies on patient types, while E. F. Long, et al. (2018) optimize bed allocation during the Ebola outbreak using greedy policies and dynamic programming.

Vaccine distribution is another critical area. Studies address both supply chain logistics (Duijzer, et al., 2018a) and allocation strategies (Tanner and Ntaimo, 2010; Yarmand, et al., 2014). Duijzer, et al. (2018b) embedd a SIR model into a vaccine distribution framework, showing that uniform allocation can reduce herd immunity effectiveness. Bertsimas, et al. (2021a) integrated fairness measures into a vaccination facility location model. The complicated transition structure leads to a non-convex problem and they use a coordinate descent algorithm to solve it. Mak, et al. (2022) design dynamic policies for two-dose COVID-19 vaccines, especially the optimal reserved level for the second does. Bennouna, et al. (2023) use a multipeak SEIRD model to guide vaccine allocation, though their approach ignored feedback effects on disease spread.

Most existing models focus on cost minimization, but the uncertain effects of these resources during rapidly evolving pandemics complicate policy design. This underscores the need for robust frameworks that account for stochasticity and evolving conditions, enabling more effective and equitable decision-making in pandemic control.

2.1.3 Emergency medicine service

The EMS problem is a critical area of operations research and healthcare management, focusing on optimizing the deployment, location, and routing of time-critical resources—such as ambulances and ventilators—to ensure timely responses to medical emergencies. Historically, the EMS problem has been addressed through deterministic models such as the location set covering problem (LSCP) and the maximal covering location problem (MCLP). These models seek to maximize coverage of demand points within a specified response time, as seen in early works like (Marianov and ReVelle, 1996). However, such deterministic approaches often fail to account for the stochastic nature of demand and travel times, leading to potential inefficiencies in real-world applications.

Motivated by these limitations, recent research has increasingly adopted probabilistic and stochastic methods to address inherent uncertainty, recognizing that EMS demand is random and travel times can vary due to traffic conditions, weather, and other factors. For instance, Beraldi, et al. (2004) proposes a stochastic programming model with probabilistic constraints for locating and sizing EMS stations. This model ensures a certain probability of coverage, addressing variability in demand and response times. Nevertheless, practical implementation is challenged by the fact that the underlying distributions are often unknown to decision-makers, necessitating estimation using historical data.

The survey X. Li, et al. (2011) highlight the use of data-driven methods to estimate demand in low- and middle-income country settings, improving the adaptability of EMS systems to dynamic conditions. When coupled with predictive tools, like machine learning, these approaches can refine estimation accuracy: however, the estimation uncertainty may compromise the quality of solutions. In this context, robust optimization serves as a suitable method to handle it. A significant study by Boutilier and T. Chan (2020) combine their prediction-optimization framework with a simulation model and real data to provide an in-depth investigation into three policy-related questions, aiming to reduce response times.

2.2 Methodologies

2.2.1 Risk optimization in markov decision processes

Since the 1950s, Markov decision processes (MDPs)—mathematical frameworks for modeling decision-making in multi-period stochastic systems, also known as stochastic dynamic programming — have been extensively studied (Bellman, 1957). Widely applied in fields such as robotics, automated control, finance, healthcare, and manufacturing (Lauri, et al., 2022; Bäuerle and Rieder, 2011; Shi, et al., 2021), MDPs traditionally focus on minimizing the expected cost, which involves finding a policy that minimizes the expected value function. However, in many applications, decision-makers are increasingly concerned not only with the expected outcome but also with the variability and risk associated with stochastic systems. This demands a framework that captures the variability a policy introduces.

To meet this need, risk-sensitive MDPs extend the classical framework by incorporating risk measures into the optimization objective. These measures include variance-related metrics (Filar, et al., 1989), percentile performance (Filar, et al., 1995), or expected exponential utility (Howard and Matheson, 1972). The primary challenge lies in formulating risk-sensitive criteria that are both conceptually meaningful and mathematically tractable. For instance, Marcus, et al. (1997) propose minimizing an exponential function of the cumulative cost, which captures not only the expected cost but also higher-order moments, while Bäuerle and Rieder (2014) generalize this approach using certainty equivalents for total or discounted costs. Recent advances include Xia (2020), who investigates infinite-stage discrete-time MDPs with a long-run average metric, incorporating both the mean and variance of rewards. Their work introduces sensitivity-based optimization theory, deriving a performance difference formula to quantify the combined mean-variance metrics of MDPs under different policies. Meanwhile, Moharrami, et al. (2022) focus on risk-sensitive exponential cost MDPs, developing a trajectory-based gradient algorithm to identify stationary points of the cost function for parameterized policies.

Within the broader field of risk-sensitive MDPs, some recent studies focus on risk-sensitive optimal stopping problems. Jelito, et al. (2021) prove the continuity of value functions for a class of continuous-time Feller-Markov processes, propose discrete-time approximations, and show their regularity. Meanwhile, Kosmala and Moriarty (2023) develop novel algorithms leveraging the geometry of functions to solve such problems. Our work in Chapter 3 contributes to this subfield by focusing on the structural properties of policies and comparing them with traditional MDP frameworks, offering new insights into the trade-offs between risk sensitivity and expected performance.

2.2.2 Robust optimization

Beyond risk measures, Robust Optimization (RO) is a powerful and widely adopted framework for decision-making under uncertainty. RO was originally developed to address the well-known *optimizer's curse* (Smith and Winkler, 2006), a phenomenon in which optimization models tend to overfit to parameter estimates derived from limited data, leading to poor out-of-sample performance. This issue is particularly prominent in stochastic programming methods relying on Sample Average Approximation (SAA), which replaces expectations or chance constraints with empirical distributions (Calafiore and Campi, 2005). While SAA is conceptually simple and asymptotically consistent under the true distribution (Shapiro, et al., 2009), it often leads to overly optimistic solutions that underestimate risk and suffer from frequent feasibility violations under realized uncertainty. Furthermore, SAA is highly sensitive to data quality; challenges such as limited sample sizes and the presence of outliers can significantly degrade solution quality and reliability.

RO mitigates these challenges by seeking decisions that hedge against worst-case deviations, thereby offering stronger feasibility guarantees. Formally, consider a cost-based attribute function $f(\boldsymbol{x}, \boldsymbol{z})$, where $\boldsymbol{x} \in \mathcal{X}$ is the *here-and-now* decision vector within the feasible set $\mathcal{X} \subseteq \mathbb{R}^D$ and $\boldsymbol{z} \in \mathcal{Z}$ is an uncertain input vector within the support set $\mathcal{Z} \subseteq \mathbb{R}^L$. In deterministic models, constraints such as $f(\boldsymbol{x}, \hat{\boldsymbol{z}}) \leq \tau$ are imposed based on a nominal input $\hat{\boldsymbol{z}} \in \mathcal{Z}$, yet they may fail under input deviations, potentially leading to violations such as $\mathbb{P}^\star \{f(\boldsymbol{x}, \hat{\boldsymbol{z}}) > \tau\}$. In contrast, robust optimization introduces the constraint

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau, \quad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma),$$

where $\mathcal{Z}(\gamma) \subseteq \mathcal{Z}$ is an uncertainty set parameterized by the budget parameter $\gamma \in [0, \infty]$, with $\mathcal{Z}(0) = \{\hat{\boldsymbol{z}}\}$ representing the deterministic case and $\mathcal{Z}(\infty) = \mathcal{Z}$ covering the full uncertainty support. Typically, $\mathcal{Z}(\gamma)$ is defined based on the discrepancy between the realized input $\boldsymbol{z}$ and the nominal input $\hat{\boldsymbol{z}}$, collecting all $\boldsymbol{z}$ such that their discrepancy remains below γ. Classical works, including Ben-Tal and Nemirovski (1998) and El Ghaoui, et al. (1998), analyze L_2-norm-based uncertainty sets, resulting in quadratic robust constraints, while Bertsimas and Sim (2004) adopt L_1-norm-based sets, leading to more tractable linear formulations.

Despite its computational appeal, RO may lead to conservative solutions if uncertainty sets are not properly calibrated. To overcome this, substantial efforts have been devoted to constructing uncertainty sets grounded in statistical principles, typically treating uncertain parameters as random variables with unknown distributions. A notable approach involves constructing uncertainty sets based on asymptotic properties derived from the central limit theorem (Bandi and Bertsimas, 2012), which provide theoretical guarantees under large samples but often suffer from conservatism in small-sample regimes, especially when data exhibit skewness or heavy tails. To alleviate conservatism, researchers have proposed uncertainty set constructions using divergence functionals (Ben-Tal, et al., 2013) and hypothesis testing methods (Bertsimas, et al., 2018). More recently, Hong, et al. (2021) developed a data-driven framework to empirically learn geometric uncertainty sets compatible with robust optimization, using data-splitting validation to tune the set's shape and size.

Building on this line of research, Chapter 5 of this thesis contributes a novel data-driven approach for jointly learning the shape and size of uncertainty sets for any conic-representable norm. Moreover, we introduce a unified framework that systematically incorporates two distinct layers of uncertainty—outcome uncertainty and estimation uncertainty—to further enhance solution robustness and stability.

2.2.3 Robust satisficing

Beyond classical robust optimization, the paradigm of Robust Satisficing (RS) has gained increasing attention in recent years as a flexible and less conservative approach to decision-making under uncertainty. The foundational concept of RS was introduced by D. B. Brown and Sim (2009), where the objective is not to optimize performance in expectation or under worst-case scenarios but to ensure that performance exceeds a specified target level under uncertainty. This shift from optimization towards target attainment naturally aligns with real-world applications where stakeholders are more concerned with meeting essential performance thresholds rather than achieving the theoretical optimum.

Ben-Tal, et al. (2017) advanced this idea by proposing the Globalized Robust Counterpart (GRC), expressed as

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa \min_{\boldsymbol{\delta} \in \mathcal{Z}(\gamma)} \Delta(\boldsymbol{z}, \boldsymbol{\delta}), \quad \forall \boldsymbol{z} \in \mathcal{Z},$$

where $\Delta(\boldsymbol{z}, \boldsymbol{\delta})$ measures the disparity between $\boldsymbol{z}$ and $\boldsymbol{\delta}$. The GRC-sum model embodies the principle of *tolerated robust satisficing*, where the optimization aims to minimize κ—the degree of tolerance—while respecting a specified target τ. This framework allows for controlled constraint violations when deviations are inevitable but penalizes the magnitude of such deviations systematically.

Extensions of this framework have appeared in various forms. D. Z. Long, et al. (2023) present a robust satisficing approach under Wasserstein ambiguity sets within distributionally robust optimization, providing analytical insights and structural results. F. Liu, et al. (2023) further extend GRC by incorporating distributional information beyond the uncertainty support, improving its modeling flexibility. These developments enable decision-makers to control not only the robustness of decisions but also the acceptable level of constraint violations as uncertainty grows. Additionally, Sim, et al. (2021) enhance the robust satisficing paradigm by integrating predictive analytics and side information, enabling data-informed control over constraint satisfaction in practical applications.

It is worth noticing that the fragility measure introduced in D. Z. Long, et al. (2023) is actually a risk measure. Also, using the Donsker-Varadhan Variational Formula (Donsker and Varadhan, 1983), we can find the dual form of entropic risk measure:

$$\frac{1}{\theta} \log \mathbb{E}_{\hat{\mathbb{P}}} \exp\left(\theta \cdot \tilde{V}^{\boldsymbol{\pi}}\right) = \sup_{\mathbb{P} \ll \hat{\mathbb{P}}} \left\{ \mathbb{E}_{\mathbb{P}}[\tilde{V}^{\boldsymbol{\pi}}] - \frac{1}{\theta} \Delta_{KL}\left(\mathbb{P} \| \hat{\mathbb{P}}\right) \right\} \leq \tau,$$

which can be viewed as a RS constraint under Kullback-Leibler (KL)-divergence.

In this thesis, we leverage the idea of RS in two works. Frist, in Chapter 4, we develop an ambiguity-aware epidemiological optimization model. Second, in Chapter 5, we integrate RS into a unified framework for robust decision models.

3 Risk-sensitive optimal stopping and applications in critical care

Recent years have witnessed a significant increase in healthcare costs, attributed to inefficiencies in various domains, including care delivery failures, uncoordinated care, overtreatment, and low-value practices (Shrank, et al., 2019). This upward trend in expenditures coincides with an increased demand for healthcare, prompting service providers, including intensive care units (ICUs), to adopt fast-track management for efficient and precise decision-making (Wilmore and Kehlet, 2001; F. Zhu, et al., 2012).

In critical care, determining when to complete an intervention is often referred to as the *stopping problem*. This concept revolves around balancing the potential benefits of ongoing support against the risks and costs—both clinical and resource-related—associated with stopping treatment. Decisions about when to stop can be influenced by a patient's evolving health status, the risk of complications, and the need to allocate resources efficiently. An example of this stopping problem is extubation for ventilated patients. Deciding when to remove a patient from mechanical ventilation is crucial because it directly affects patient safety, recovery, and overall outcomes. Extubating too early raises the risk of extubation failure, leading to re-intubation and increased mortality (Epstein, et al., 1997; Thille, et al., 2013). In contrast, prolonged intubation can result in complications such as ventilator-associated pneumonia, airway trauma, and respiratory muscle fatigue. Similarly, determining patient discharge timing is tricky. Premature discharge can increase readmission, and delayed discharge can waste resources. In extubation and discharge decisions, the goal is to find the balance that safeguards patient well-being, enhances outcomes, and uses resources responsibly.

To analyze the trade-off, two main performance metrics play an important role: the probability of retreatment and the associated costs. The probability of retreatment, or failure rate, indicates the proportion of patients requiring re-treatment; lower rates reflect more effective decisions and improved patient outcomes. The cost of retreatment, often quantified by the LOS in the ICU, reflects the time a patient needs under care, with shorter stays indicating more efficient resource utilization. Traditionally, expected LOS is commonly used to evaluate the outcomes of stopping decisions. However, relying solely on expected values fails to account for the inherent risks and variability associated with stopping treatments.

Example 3.1. *For example, in the context of the discharge problem, the terminal risk can be measured by* $\tilde{g} = \tilde{r}\tilde{l}$*, where* $\tilde{r}$ *represents the readmission event following Bernoulli distribution, and* $\tilde{l}$ *is the LOS during readmission, usually modeled by a log-normal distribution. Assuming* $\tilde{r} \sim \text{Bernoulli}(0.1)$ *and* $\tilde{l} \sim \text{LogNormal}(1, 2^2)$*, the expected cost,* $\mathbb{E}_{\mathbb{P}}[\tilde{g}] = 2.01$ *days, appears low. However, this expected value obscures significant risks:* 2.6% *patients experience readmissions lasting more than 10 days, and* 5% *CVaR is approximately* 39.25 *days.*

These prolonged ICU stays due to readmissions not only deplete resources but also lead to health complications and potential patient mortality (Joynt, Jha, et al., 2012; Zuckerman, et al., 2016). This example underscores why it is crucial to incorporate risk metrics into decision-making processes, as reliance solely on expected values could prompt decision-makers to adopt overly aggressive discharge strategies.

In this research, we explore two main questions:

- **Whether and how risks in medical treatment stopping problems should be measured?**
- **How might incorporating a risk measure influence the development of optimal treatment policies?**

To address these questions, it is crucial to adopt a metric that accounts for the risks associated with uncertain outcomes, rather than relying solely on expectations, to effectively evaluate performance. In this study, we focus on the entropic risk measure, which offers unique properties for capturing risk under uncertainty and ensures the tractability of the optimal stopping problem. Embedded with the entropic risk measure, the decision-making process can be formulated as a risk-sensitive MDP, or more specifically, a risk-sensitive optimal stopping problem. Following the theoretical developments, we undertake an analysis of the structural properties of the value function and optimal policy. Subsequently, we carry out two case studies to empirically evaluate the performance of our model.

The contributions of this research are as follows: [1]

1. Building upon the risk-sensitive MDP framework, which comprehensively incorporates risk elements and high-order-moment information, our models adeptly quantify the uncertainty

[1]This chapter is based on the working paper: Lou, Z., and Xie, J. (2024). Mitigating Decision Risks in Critical Care. Co-authorship details: Prof. Dr. Jingui Xie (Affiliation: TUM) supervised the research design, advised on the entropic risk measure formulation, and reviewed theoretical proofs. Z. Lou (PhD candidate) developed the risk-sensitive MDP framework, conducted ICU extubation/discharge experiments, drafted the chapter, and revised it based on Prof. Xie's feedback. Both authors approved the final chapter.

inherent in both state transitions and termination costs. A thorough investigation into the structural properties of the value function and policy is undertaken. We prove the submodularity of the value function under certain assumptions, which confirms the existence of a threshold structure in the optimal strategy, a critical aspect for practical implementation.

2. A comparative study is conducted between our model and the nominal MDP approach. This examination reveals specific circumstances under which our model adopts more aggressive/conservative strategies compared to nominal MDP, and elucidates a nuanced trade-off between termination and continuation risks in medical decision-making.

3. We also incorporate predicted future states into our model. Under certain assumptions, we establish that the resulting optimal policy maintains a threshold structure.

4. To support our theoretical developments, we conduct several numerical experiments on extubation and discharge decisions in ICUs. These experiments demonstrate that our model delivers an improved risk profile, albeit at the expense of some trade-off in expected performance.

Organization: The rest of this chapter is structured as follows: We introduce the risk measure in §3.1. The formulation of the stopping decision problem is detailed in §3.2. We compare the risk-sensitive optimal stopping policies and risk-neutral ones in §3.3. We extend our model with predictive information in §3.4. Numerical experiments are discussed in §3.5, and the paper concludes with §3.6. All the technical proofs are provided in Appendix A.1.

3.1 Metrics for evaluating decision risk

In this section, we discuss how to measure the decision risk associated with stopping treatment. To evaluate the uncertain outcomes associated with stopping decisions, specifically the probability of re-treatment, along with the associated cost of re-treatment, we aim to find a risk measure that can capture the outcome risk and can be effectively integrated into dynamic decision-making.

Traditionally, the expected LOS has been a key metric for assessing the outcomes of medical interventions. Fast-track protocols, such as early extubation or discharge, encourage healthcare providers to minimize ICU LOS (Wong, et al., 2016). Additionally, expected mortality rates help evaluate service quality (Moran, et al., 2008). These expected metrics are widely used in dynamic decision-making processes, as the nominal MDP keeps good properties for prescriptive analysis. However, relying solely on expected values overlooks inherent risks.

3.1.1 Stopping risk metrics

To address the above issues, we seek a metric that can capture higher-order statistical information and tail distribution, while retaining some properties convenient for downstream prescriptive optimization models, such as computational tractability. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathcal{V}$ denote the set of random variables on this probability space. The intended stopping risk metric should be a functional from $\mathcal{V}$ to $\mathbb{R} \cup \{+\infty\}$ that is equipped with a hyper-parameter reflecting the risk preference of decision makers and satisfies the following properties. Among these properties, monotonicity aligns with our goal of mitigating the risk of high-cost outcomes, while maintaining tractability. Translation invariance ensures that as θ increases, only the risk component increases, leaving the deterministic (risk-neutral) term unaltered.

Definition 3.1 (Stopping Risk Metric.)**.** *The functional* $\mu_\theta : \mathcal{V} \to \mathbb{R} \cup \{+\infty\}$ *is a* stopping risk metric *if and only if* μ_θ *is lower semi-continuous and for all* $\tilde{v}, \tilde{v}_1, \tilde{v}_2 \in \mathcal{V}$ *it satisfies:*

- *(Risk-neutral)* $\mu_0(\tilde{v}) = \mathbb{E}_{\mathbb{P}}[\tilde{v}]$.
- *(Monotonicity)* $\mu_\theta(\tilde{v})$ *is increasing in* θ. *If* $\tilde{v}_1 \geq_{st} \tilde{v}_2$, *then* $\mu_\theta(\tilde{v}_1) \geq \mu_\theta(\tilde{v}_2)$ *for any* $\theta \geq 0$.
- *(Normalization)* $\mu_\theta(0) = 0$, *for any* $\theta \geq 0$.
- *(Translation invariance) For any* $\tilde{v} \in \mathcal{V}, c \in \mathbb{R}, \theta \geq 0$, $\mu_\theta(\tilde{v} + c) = \mu_\theta(\tilde{v}) + c$.

Here, the hyper-parameter $\theta \in [0, +\infty)$ serves as an indicator of the decision maker's risk tolerance. When θ is larger, the decision makers are more risk aversive. In the case where $\theta = 0$, they are risk neutral, and the risk measure aligns with the expected outcome. In the context of healthcare, monotonicity implies that when a patient is almost surely in a more dangerous status than the other, his(her) stopping risk metric is consistently higher. Moreover, the translation invariance guarantees that the value a risk-free component contributes to the risk metric does not change with respect to θ, such as the fixed inspection fee.

When choosing an appropriate metric, one critical factor in selecting an appropriate metric is its ability to preserve key structural properties of the dynamic decision-making process. For example, risk measures such as Value at Risk (VaR) or Conditional Value at Risk (CVaR) result in optimal policies that are history-dependent in sequential decision-making (Bäuerle and Ott, 2011; X. Li, et al., 2022). While these measures are valuable for capturing risk, they result in non-Markovian policies that complicate the design of algorithms and increase computational complexity. Additionally, history-dependent optimal policies can obscure management insights, making it challenging to communicate actionable strategies to physicians. In contrast, selecting a metric that retains the Markovian property would not only simplify the algorithm design process but also facilitate clearer, more interpretable decision-making frameworks in clinical or operational settings.

3.1.2 Entropic risk measure

Given the above considerations, we propose using the entropic risk measure to evaluate the risk of the cumulative cost. This risk measure is widely used in stochastic optimization problems (see, e.g., Xie, et al., 2023; T. Zhu, et al., 2023). Unlike VaR and CVaR, which focus exclusively on the tail of the distribution, the entropic risk measure accounts for the entire distribution of outcomes. More importantly, the entropic risk measure offers key advantages, such as existing an optimal deterministic Markovian policy (Hau, et al., 2023) and ensuring submodularity, which will be elaborated upon in the subsequent sections.

The main idea behind the entropic risk measure is to quantify risk using the principles of entropy, a concept from thermodynamics and information theory. Let $\mathcal{V}$ be the space of random variables. The definition is as follows.

Definition 3.2. ***(Entropic Risk Measure)*** *Given a random variable $\tilde{v} \sim \mathbb{P}$, the entropic risk measure is defined as the functional $\mu : \mathcal{V} \to \mathbb{R}$,*

$$\mu_\theta(\tilde{v}) = \begin{cases} \dfrac{1}{\theta} \, log \, \mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v} \cdot \theta\right)\right] & if\ \theta > 0, \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{v}\right] & if\ \theta = 0. \end{cases}$$

The entropic risk measure satisfies all conditions for the stopping risk metric previously discussed. Firstly, we show that the entropic risk measure satisfies the properties in Definition 3.1:

- (Risk-neutral) $\mu_0(\tilde{v}) = \mathbb{E}_{\mathbb{P}}\left[\tilde{v}\right]$.
- (Monotonicity) From Jensen's inequality,

  $$\left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_1 \tilde{v}\right)\right]\right]^{\theta_2/\theta_1} \leq \mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_2 \tilde{v}\right)\right],$$

  which implies that $\left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_1 \tilde{v}\right)\right]\right]^{1/\theta_1} \leq \left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_2 \tilde{v}\right)\right]\right]^{1/\theta_2}$ for any $\theta_1 \leq \theta_2$. Therefore, the statement w.r.t. θ holds. Given the condition that $\theta > 0$, the function $\exp(\theta x)$ is increasing w.r.t. x. Then, according to the property of First-order stochastic dominance (FSD), we have

  $$\mathbb{E}_{\mathbb{P}}\left[\exp(\theta \tilde{v}_1)\right] \geq \mathbb{E}_{\mathbb{P}}\left[\exp(\theta \tilde{v}_2)\right],$$

  leading to the inference that $\mu_\theta(\tilde{v}_1) \geq \mu_\theta(\tilde{v}_2)$. In the case where $\theta = 0$, the inequality $\mathbb{E}_{\mathbb{P}}\left[\tilde{v}_1\right] \geq \mathbb{E}_{\mathbb{P}}\left[\tilde{v}_2\right]$ can be derived again, leveraging the property of FSD.
- (Normalization) $\mu_\theta(0) = \frac{1}{\theta} \log \mathbb{E}_{\mathbb{P}}\left[e^0\right] = 0$.
- (Translation invariance)

  $$\mu_\theta(\tilde{v} + c) = \frac{1}{\theta} \log \mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta \cdot (\tilde{v} + c)\right)\right] = \frac{1}{\theta} \log\left(\exp\left(\theta c\right) \mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta \cdot \tilde{v}\right)\right]\right) = \mu_\theta(\tilde{v}) + c.$$

The properties of normalization and translation invariance follow directly from the definition of $\mu_\theta(\tilde{v})$.

Moreover, we can prove that $(\theta, \tilde{v}) \mapsto \mu_{\frac{1}{\theta}}(\tilde{v})$ is jointly convex in $(\theta, \tilde{v})$:

Using Hölder's inequality, we can show that $\mathbb{E}_{\mathbb{P}}\left[\exp\left(\lambda \tilde{v}_1\right) \cdot \exp\left((1-\lambda)\tilde{v}_2\right)\right] \leq \left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v}_1\right)\right]\right]^{\lambda} \cdot \left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v}_2\right)\right]\right]^{1-\lambda}$. This implies that

$$\log \mathbb{E}_{\mathbb{P}}\left[\exp\left(\lambda \tilde{v}_1 + (1-\lambda)\tilde{v}_2\right)\right] \leq \lambda log\, \mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v}_1\right)\right] + (1-\lambda) log\, \mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v}_2\right)\right],$$

i.e., $log\, \mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v}\right)\right]$ is convex in $\tilde{v}$. The perspective of $log\, \mathbb{E}_{\mathbb{P}}\left[\exp\left(\tilde{v}\right)\right]$, that is $\mu_{\frac{1}{\theta}}(\tilde{v})$, preserves the convexity (the discussion of perspective operation can be found in §3.2.6 in Boyd and Vandenberghe (2004)). This completes the proof that the entropic risk measure satisfies the definition of a stopping risk metric.

As the parameter θ increases, the decision makers become more risk aversive. It converges to the essential supremum of $\tilde{v}$ as θ increases to $+\infty$. If $\theta = 0$, it is equivalent to the expectation. Also, the entropic risk measures have some desirable properties in violation probability and expected shortages. In the following proposition, we demonstrate that the violation probability and expected shortage are upper bounded by functions of θ, given that $\mu_\theta(\tilde{v}) \leq \tau$.

Proposition 3.1. *Suppose $\mu_\theta(\tilde{v}) \leq \tau$. For all $\epsilon \geq 0$, we have*

$$\mathbb{P}\left[\tilde{v} > \tau + \epsilon\right] \leq \exp\left(-\epsilon\theta\right), \tag{3.1}$$

and

$$\mathbb{E}\left[(\tilde{v} - \tau - \epsilon)^{+}\right] \leq \frac{1}{\theta e} \exp\left(-\epsilon\theta\right). \tag{3.2}$$

Proof. The first inequality is a quick application of Chernoff bound. The second one is due to the fact that $x^{+} \leq \exp\left(x - 1\right)$ and $\mathbb{E}\left[(\tilde{z} - \tau - \epsilon)^{+}\right] \leq \frac{1}{\theta e}\mathbb{E}[\exp\left((\tilde{z} - \tau - \epsilon)\theta\right)] \leq \frac{1}{\theta e} \exp\left(-\epsilon\theta\right)$. □

Employing the entropic risk measure allows us to effectively capture the risk while also accounting for expected performance. In practice, determining an appropriate value for θ is challenging. One potential approach is to leverage prior knowledge, such as the magnitude of standard deviation. Alternatively, we suggest the satisficing approach for selecting θ (D. B. Brown and Sim, 2009). Given a target τ, for instance $\tau = k\mathbb{E}_{\mathbb{P}}\left[\tilde{v}\right]$, we find the maximal θ such that $\mu_\theta(\tilde{v}) \leq \tau$. It aims to find a decision that meets the target as much as possible. This procedure can be done via Bisection search (see Appendix A.4, which also includes a numerical example).

3.2 Model formulation

In this section, we present the optimal stopping problem for medical treatment and explore the concept of the stopping risk metric within the MDP framework. We begin by detailing the fundamental components of the MDP model.

Time horizon. We assume that decisions are made within discrete-time epochs indexed by $t \in \mathcal{T}$. We consider a finite-time horizon MDP, *i.e.*, $\mathcal{T} = \{0, \ldots, T\}$.

States. To capture patients' health condition, we classify patients into one of $\mathcal{S}$ based on the observed clinical variables, with each class $s \in \mathcal{S} = \{1, 2, \ldots, S\}$ potentially corresponding to a particular health condition. Class 1 indicates the most severe condition, while class S represents the most favorable health state. The notation s_t is used to signify the patients' class at any epoch t, with s_0 marking the initial state. The initial state follows a distribution $\boldsymbol{p}_0$.

Actions. In each decision epoch, the Decision maker (DM) has to decide whether to stop the treatment or not. The action in epoch $t \in \mathcal{T}$ is $a_t \in \mathcal{A} = \{0, 1\}$, where 0 indicates continuation and 1 indicates stopping. We assume the treatment should be stopped despite the patient's health condition in the last epoch, *i.e.*, $a_T = 1$.

Transitions. If the DM decides to continue in epoch t, *i.e.*, $a_t = 0$, the patient in class i will evolve to class j in epoch $t + 1$ with probability $P(i, j)$, where $i, j \in \mathcal{S}$. Otherwise, the decision process ends if the treatment is terminated, *i.e.*, $a_t = 1$. Let $\mathbf{P}$ represent the transition matrix. $\boldsymbol{P}_i$ denotes the i-th row, which encapsulates the probability mass function governing the transition dynamics to the subsequent state s_{t+1} given the current state $s_t = i$.

Costs. In each decision epoch, the corresponding cost depends on the patient's current class and the action taken by the DM. If the DM decides to continue the treatment, the immediate cost is a deterministic function $H(s)$ depending on the patient's state. If the DM decides to stop the treatment, there is a random terminal cost $\tilde{G}(s)$.

The decision-making process involves sequential decisions at each epoch, where the decision-maker must evaluate the patient's condition, represented by the current state, to determine whether to stop the treatment. Once the process concludes, the risks and costs of re-treatment are reflected in the terminal cost.

Policy. We define the policy sequence as $\boldsymbol{\pi} = (\pi_0, \pi_1, \cdots, \pi_T)$, where π_t specifies the decision rule at each time step t. Let Π_{MD} denote the set of all Markovian deterministic policies, where each policy π_t is a mapping function $\pi_t : \mathcal{S} \to \mathcal{A}$. Similarly, let Π_{HR} represent the set of all history-dependent randomized policies, where each π_t is defined as a mapping $\pi_t : \mathcal{S}^{t+1} \to \mathcal{P}(\mathcal{A})$, with $\mathcal{P}(\mathcal{A})$ as the set of all distributions over $\mathcal{A}$. Then, we introduce an order relation on the set of Markovian deterministic policies as follows:

Definition 3.3. *Consider two Markovian deterministic policies, $\boldsymbol{\pi}$ and $\boldsymbol{\pi}^\dagger \in \Pi_{\mathrm{MD}}$.*

- *We say that* $\boldsymbol{\pi}$ *represents a more proactive stopping decision compared to* $\boldsymbol{\pi}^\dagger$ *if* $\{s \in \mathcal{S} \,|\, \pi_t^\dagger(s) = 1\} \subseteq \{s \in \mathcal{S} \,|\, \pi_t(s) = 1\},\ \forall t \in \mathcal{T}$.

- *Conversely, we say that* $\boldsymbol{\pi}$ *represents a more conservative stopping decision compared to* $\boldsymbol{\pi}^\dagger$ *if* $\{s \in \mathcal{S} \,|\, \pi_t^\dagger(s) = 0\} \subseteq \{s \in \mathcal{S} \,|\, \pi_t(s) = 0\},\ \forall t \in \mathcal{T}$.

3.2.1 Risk-sensitive optimal stopping

Equipped with the entropic risk measure, our Risk-sensitive optimal stopping problem (ROSP) is formulated as

$$\min_{\boldsymbol{\pi} \in \Pi_{\mathrm{H}R}} \ \mu_\theta \left(\tilde{V}^{\boldsymbol{\pi}}(s) \right). \tag{3.3}$$

Note that the Nominal optimal stopping problem (NOSP) is a special case of ROSP when $\theta = 0$. ROSP ensures the existence of a Markov deterministic optimal policy. Specifically, as shown by Hau, et al. (2023), there exists a $\boldsymbol{\pi}_\theta^\star \in \Pi_{\mathrm{M}D}$ such that:

$$\mu_\theta \left(\tilde{V}^{\boldsymbol{\pi}_\theta^\star}(s) \right) = \min_{\boldsymbol{\pi} \in \Pi_{\mathrm{H}R}} \ \mu_\theta \left(\tilde{V}^{\boldsymbol{\pi}}(s) \right).$$

Therefore, our analysis will focus on deterministic Markovian policies.

The ROSP framework adopts a risk measure for the total cost as its optimization target, inherently prioritizing a better risk profile over purely optimizing expected performance. This trade-off results in a certain degree of sacrifice in the expected value of the cost to achieve enhanced risk management. Mathematically,

$$\mathbb{E}_{\mathbb{P}} \left[\tilde{V}^{\boldsymbol{\pi}_\theta^\star}(s) \right] \geq \min_{\boldsymbol{\pi} \in \Pi_{\mathrm{H}R}} \ \mathbb{E}_{\mathbb{P}} \left[\tilde{V}^{\boldsymbol{\pi}}(s) \right].$$

In most scenarios, the inequality is strict.

The terminal cost $\tilde{G}$ exhibits significant variability of therapy outcome, largely attributable to the inherent risk of terminating treatment prematurely. For instance, in the context of extubation, a failed extubation attempt can exacerbate a patient's condition, leading to reintubation and an extended duration of hospitalization. To accommodate such risks, our model incorporates a stochastic terminal cost. Some examples of terminal distributions and the corresponding risk measures are presented in Appendix A.3. For the subsequent analysis, we assume that higher order moments of $\tilde{G}(s)$ exist for every state $s \in \mathcal{S}$. This ensures that the function $\mu_\theta \left(\tilde{G}(s) \right)$ is well-defined and exhibits smooth behavior around $\theta = 0$. By adopting this assumption, we circumvent the degenerate case where $\mu_\theta(\cdot)$ becomes $+\infty$. Additionally, we impose the following assumption regarding the monotonicity of $\mu_\theta \left(\tilde{G}(s) \right)$.

Assumption 3.1. *For any pair of states s_1 and s_2 where $s_1 \geq s_2$, it holds that $\tilde{G}(s_1) \leq_{st} \tilde{G}(s_2)$. This implies that the function $\mu_\theta\left(\tilde{G}(s)\right)$ decreases with the state s for any $\theta \geq 0$. It also holds that $H(s_1) \leq H(s_2)$.*

Assumption 3.1 suggests that patients in better health conditions are likely to have a smaller cost after stopping treatment. This correlation is indicative of either a quicker recovery time or a reduced likelihood of extubation failure. Additionally, this assumption posits that the continuation costs for patients in better conditions do not surpass those in worse conditions.

Let the U-function $U_t^{\boldsymbol{\pi}}(\theta, s) := \mu_\theta\left(\tilde{V}_t^{\boldsymbol{\pi}}(s)\right)$ denote the risk-to-go conditional on state s at time t. Here, $\tilde{V}_t^{\boldsymbol{\pi}}(s) = \sum_{\tau=t}^{T} \tilde{c}_\tau\left(s_\tau, \pi_\tau(s_\tau)\right) \big| s_t = s$, where $\tilde{c}_\tau\left(s_\tau, \pi_\tau(s_\tau)\right)$ denotes the one-period cost of taking action $\pi_\tau(s_\tau)$. Based on the above setups, the U-function has the following backward recursive equation:

$$U_t^{\boldsymbol{\pi}}(\theta, s) = \begin{cases} \mu_\theta\left(\tilde{G}(s)\right) & \text{if } \pi_t(s) = 1, \\ H(s) + \dfrac{1}{\theta} \log \displaystyle\sum_{s'} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}^{\boldsymbol{\pi}}(\theta, s')\right) & \text{if } \pi_t(s) = 0. \end{cases}$$

The optimal U-function is defined as

$$U_0(\theta, s) = \min_{\boldsymbol{\pi}} U_0^{\boldsymbol{\pi}}(\theta, s). \tag{3.4}$$

It is known that an optimal policy exists and can be obtained via a sequence of backward minimization problems.

Proposition 3.2 (Bäuerle and Rieder (2014))**.** *Given $\theta > 0$, the optimal policy*

$$\boldsymbol{\pi}_\theta^* = (\pi_{\theta,0}^*, \pi_{\theta,1}^*, \cdots, \pi_{\theta,T}^*)$$

for the finite horizon ROSP is determined through a backward recursion as follows: Begin with $U_T(\theta, s) = \mu_\theta(\tilde{G}(s))$. Then, for $t = T-1, \cdots, 0$, compute

$$U_t(\theta, s) = \min\left\{\mu_\theta\left(\tilde{G}(s)\right), H(s) + \frac{1}{\theta} \log \sum_{s' \in \mathcal{S}} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}(\theta, s')\right)\right\}; \tag{3.5}$$

and the optimal policy is given by

$$\pi_{\theta,t}^\star(s) = \begin{cases} 1 & \textit{if } \mu_\theta\left(\tilde{G}(s)\right) \leq H(s) + \dfrac{1}{\theta} \log \displaystyle\sum_{s' \in \mathcal{S}} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}(\theta, s')\right), \\ 0 & \textit{else.} \end{cases} \tag{3.6}$$

Furthermore, the optimal risk-to-go $\min_{\boldsymbol{\pi}} U_t^{\boldsymbol{\pi}}(\theta, s)$ satisfies Equation (3.5).

Remark 3.1. *For any initial state s_0 following the distribution $p_0(\cdot)$, the risk measure of cumulative cost under $\boldsymbol{\pi}^*$ is calculated by $\mu_\theta\left(\tilde{V}^{\boldsymbol{\pi}^\star}\right) = \frac{1}{\theta}\log\left(\sum_{s\in\mathcal{S}} p_0(s)\cdot\exp\left(\theta\cdot U_0(\theta,s)\right)\right)$. For notation convenience and clarity in our analysis, we denote the optimal policy sequence associated with a given risk tolerance level θ as $\boldsymbol{\pi}_\theta^\star$.*

Proposition 3.2 asserts that on any optimal trajectory, the subsequent path must be the optimal solution for the subproblem starting at that juncture. In essence, for an optimal policy $\boldsymbol{\pi}^\star$ obtained from Equation (3.6), the subset $\{\pi_t^*, ..., \pi_T^*\}$ minimizes the subproblem $\min_{\boldsymbol{\pi}} \mu_\theta\left(\tilde{V}_t^\pi(s)\right)$ for every $s\in\mathcal{S}$ and $t\in\mathcal{T}$.

Building upon the recursion of U-function, the subsequent proposition delves into a detailed investigation of the monotonic properties of $U_t(\theta,s)$.

Proposition 3.3. *The U-function satisfies the following properties:*

(a) $U_t(\theta,s)$ is increasing in both t and θ.

(b) If $\boldsymbol{P}_i \leq_{st} \boldsymbol{P}_j$, $\forall i\leq j\in\mathcal{S}$, then, under Assumption 3.1, $U_t(\theta,s)$ is decreasing in s for any $t\in\mathcal{T}$ and $\theta\in\mathbb{R}_+$.

In Proposition 3.3, it is established that the U-function exhibits an increasing trend with respect to t, attributable to the termination effect. Additionally, it demonstrates an increasing pattern in relation to θ, reflecting the inherent risk-averse characteristic. FSD condition ensures that patients in better health are more likely to progress to a better health status. Together with Assumption 3.1, we deduce that a healthier patient has a lower risk value.

In the following example, we compare the distributions of total cost from NOSP and ROSP by using the Monte Carlo simulation with 10^4 samples.

Example 3.2. *Consider a scenario to demonstrate the balance between risk tolerance and expected cost. Assume two states $\mathcal{S}=\{1,2\}$ and three periods $\mathcal{T}=\{0,1,2\}$. DM must stop treatment for all patients at period 2. The state transition probabilities are given by the matrix $\boldsymbol{P}=\begin{pmatrix}0.7 & 0.3; 0.1 & 0.9\end{pmatrix}$. The terminal cost for state 1 follows $\tilde{G}(1)\sim\mathcal{N}(130,20)$ and for state 2, $\tilde{G}(2)\sim\mathcal{N}(50,5)$. The fixed immediate continuation cost is 20. The initial state s_0 is 1. We compare the performance of NOSP and ROSP with $\theta=0.1$. Figure 3.1 shows the distributions of $\tilde{V}^{\boldsymbol{\pi}_0^\star}$ and $\tilde{V}^{\boldsymbol{\pi}_\theta^\star}$. The key difference between NOSP and ROSP policies occurs at period 1: NOSP only terminates treatment for those in state 2, while ROSP terminates all patients. Figure 3.1 presents empirical distributions of $\tilde{V}^{\boldsymbol{\pi}_0^\star}$ and $\tilde{V}^{\boldsymbol{\pi}_\theta^\star}$. ROSP achieves a better risk profile and a lighter tail in cost distribution, though it comes with a slightly higher expected cost. Specifically, $\mathbb{E}_\mathbb{P}\left[\tilde{V}^{\boldsymbol{\pi}_0^\star}\right]=123.20$ and $\mathbb{E}_\mathbb{P}\left[\tilde{V}^{\boldsymbol{\pi}_\theta^\star}\right]=130.00$. In terms of risk, $\mathrm{VaR}_{5\%}\left(\tilde{V}^{\boldsymbol{\pi}_\theta^\star}\right)=162.89$ outperforms $\mathrm{VaR}_{5\%}\left(\tilde{V}^{\boldsymbol{\pi}_0^\star}\right)=195.39$.*

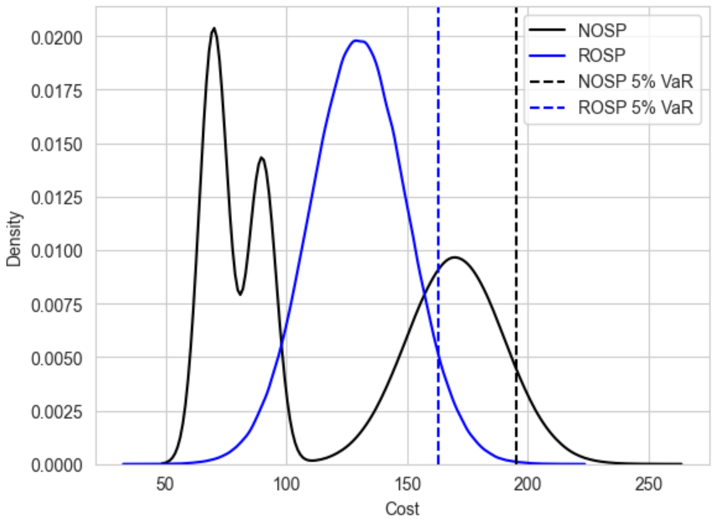


Figure 3.1: Distributions of total costs from NOSP and ROSP

3.3 Optimal risk-sensitive stopping policy

In this section, we examine the structural properties of the optimal risk-sensitive stopping policy. To understand how risk considerations influence decisions, we then compare the policies derived from ROSP and NOSP to identify conditions under which the risk-sensitive approach leads to more conservative or proactive policies.

First, to analyze the optimal policy, we define a function that captures the difference between the risk of stopping and the risk of continuing treatment, corresponding to the two terms on the right-hand side of Equation (3.5).

Definition 3.4. *(***Benefit Function***) We define the benefit function of continuing the treatment as $h_t^\theta(s)$, where*

$$h_t^\theta(s) = \begin{cases} \mu_\theta\left(\tilde{G}(s)\right) - H(s) - \dfrac{1}{\theta}\log \displaystyle\sum_{s' \in \mathcal{S}} P(s,s')\exp\left(\theta \cdot U_{t+1}(\theta, s')\right) & \text{if } \theta > 0, \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s)\right] - H(s) - \displaystyle\sum_{s' \in \mathcal{S}} P(s,s')U_{t+1}(0,s') & \text{if } \theta = 0. \end{cases} \tag{3.7}$$

Intuitively, it is optimal to terminate the treatment if $h_t^\theta(s) \leq 0$, which means there is no benefit of continuing treatment. Building upon this understanding, the subsequent analysis will explore the monotonicity of $h_t^\theta(s)$ in time index t.

Proposition 3.4. *The benefit $h_t^\theta(s)$ of continuing the treatment is decreasing in time t for all $s \in \mathcal{S}$ and $\theta \in \mathbb{R}_+$.*

This Proposition suggests that as time t increases, there are more states where the benefit $h_t^\theta(s)$ becomes negative. Importantly, the set $\{s \in \mathcal{S} \mid h_t^\theta(s) \leq 0\}$ represents the collection of states where the optimal policy dictates termination. Therefore, as time progresses, our model tends to adopt an increasingly "aggressive" stopping policy.

3.3.1 Structure of U-function

To derive the structural results, we rely on the following assumptions. We first introduce the concept Monotone likelihood ratio (MLR) dominance, and then Assumption 3.2.

Definition 3.5. *(**Monotone Likelihood Ratio**) A probability vector $\boldsymbol{P}_j$ dominates $\boldsymbol{P}_i$ in MLR ordering if $P(j,s_2)P(i,s_1) \geq P(i,s_2)P(j,s_1)$, for $s_1 < s_2$, represented as $\boldsymbol{P}_j \geq_{lr} \boldsymbol{P}_i$.*

The MLR order represents a stronger condition than FSD, such that $\boldsymbol{P}_j \geq_{lr} \boldsymbol{P}_i$ implies $\boldsymbol{P}_j \geq_{st} \boldsymbol{P}_i$. The concept of MLR dominance is widely utilized in dynamic programming, particularly in partially observable MDPs. This is because it is maintained under conditional expectations and supports the establishment of a threshold policy structure (Krishnamurthy, et al., 2018). In the context of healthcare, this means that for a patient with a better health status (represented by a larger s), the probability of observing improved health outcomes is consistently higher compared to a patient in a more severe condition.

Assumption 3.2. *Transition matrix $\mathbf{P}$ is Total positivity of order 2 (TP2). Equivalently, $\boldsymbol{P}_i \leq_{lr} \boldsymbol{P}_j$ for any $i < j$. (Krishnamurthy, 2016)*

Proposition 3.5. *Under Assumption 3.2, assume that f is submodular over $\{t, t+1\} \times \mathcal{S}$, i.e., for any $s_1 < s_2 \in \mathcal{S}$, the inequality $f_{t+1}(s_1) - f_{t+1}(s_2) \geq f_t(s_1) - f_t(s_2)$ holds. Then, for any $\tilde{z}_i \sim \boldsymbol{P}_i$ and $\tilde{z}_j \sim \boldsymbol{P}_j$ ($i < j \in \mathcal{S}$), and $\theta > 0$,*

$$\mu_\theta\left(f_{t+1}(\tilde{z}_i)\right) - \mu_\theta\left(f_{t+1}(\tilde{z}_j)\right) \geq \mu_\theta\left(f_t(\tilde{z}_i)\right) - \mu_\theta\left(f_t(\tilde{z}_j)\right). \tag{3.8}$$

Proposition 3.5 reveals a significant link between submodular functions and the entropic risk measure in the context of TP2 transitions. It establishes that the entropic risk measure applied to random variables under TP2 transitions retains the submodular structure of the value function. This insight is crucial in supporting the assertions made in Theorem 3.1.

Assumption 3.3. *The function $h^\theta_{T-1}(s)$ is decreasing in $s \in \mathcal{S}$.*

Note that $h^\theta_{T-1}(s) = \mu_\theta\left(\tilde{G}(s)\right) - H(s) - \frac{1}{\theta} \cdot \log \sum_{s'} P(s,s') \cdot \exp\left(\theta \cdot \mu_\theta\left(\tilde{G}(s')\right)\right)$. This assumption implies that as a patient's health status improves (*i.e.*, as s increases), the incremental benefit from an extra period of treatment prior to stopping decreases. It ensures that for any $s_1 < s_2 \in \mathcal{S}$,

$$\max\{0, h^\theta_{T-1}(s_1)\} = U_T(\theta, s_1) - U_{T-1}(\theta, s_1) \geq U_T(\theta, s_2) - U_{T-1}(\theta, s_2) = \max\{0, h^\theta_{T-1}(s_2)\},$$

thus affirming submodularity at the beginning of backward recursion. With these assumptions in place, we can proceed to elucidate the characteristics of the optimal policy's structure.

Theorem 3.1 (Optimal Threshold Policy)**.** *Under Assumptions 3.2 and 3.3, we have:*

1. *There exists an optimal threshold $s^{\star}_{t,\theta}$ such that we stop the treatment if $s_t \in \mathcal{R}^1_{t,\theta} = \{s \geq s^{\star}_{t,\theta}\}$ and continue the treatment if $s_t \in \mathcal{R}^0_{t,\theta} = \{s < s^{\star}_{t,\theta}\}$ at epoch $t \in \mathcal{T}$. In addition, for all $t \in \mathcal{T} \backslash \{T\}$, we have $\mathcal{R}^0_{t+1,\theta} \subset \mathcal{R}^0_{t,\theta}$ and $\mathcal{R}^1_{t,\theta} \subset \mathcal{R}^1_{t+1,\theta}$.*

2. *$U_t(\theta, s)$ is submodular in $\mathcal{T} \times \mathcal{S}$.*

Theorem 3.1 establishes the existence of a threshold structure within the optimal policy of ROSP. It delineates that for any given $\theta > 0$, the designated termination region $\mathcal{R}^1_{t,\theta}$ progressively expands with the increase in time t. Furthermore, the property of an optimal threshold is, in fact, equivalent to the condition that for any $t < T$, the utility function $U_t(\theta, s)$ is submodular over $\{t, T\} \times \mathcal{S}$. The second point of the theorem asserts a more potent result, indicating that this submodularity extends over the entire space $\mathcal{T} \times \mathcal{S}$. This fundamental insight into the submodularity property provides a cornerstone for understanding the structural attributes of the U-function.

Remark 3.2. *While part 1 of Theorem 3.1 has a similar structure as Theorem EC.1 in Cheng, et al. (2024), the proof of Theorem 3.1 follows a distinct approach. The primary challenge arises from the fact that the exponential component is not linearly separable. Unlike NOSP, which relies on the assumption of first-order stochastic dominance, our results require a stronger condition—specifically, the TP2 property or, at a minimum, the reverse hazard rate order.*

Remark 3.3. *Although Assumptions 3.2 and 3.3 provide sufficient conditions for the theorem's conclusions, they are not strictly necessary. Notably, the threshold structure persists even in cases where the transition matrix does not satisfy TP2. For example, in the extubation decision problem (Section 3.5.1), the transition matrix* (3.12) *does not exhibit the TP2 property, yet the threshold structure remains evident, as illustrated in Figure 3.2.*

3.3.2 ROSP v.s. NOSP

In our following analysis, we theoretically contrast the optimal policies derived from the ROSP and the nominal optimal stopping models. The divergence in these policies mainly stems from how they balance the uncertainty in terminal costs against the unpredictability in the progression of health conditions.

We begin by examining a scenario where the terminal cost is deterministic.

Theorem 3.2 (Proactive Policy)**.** *In cases where the terminal cost is deterministic, signified by $G(s)$, the ROSP model's optimal treatment policy will be more proactive in stopping treatment* *compared to that of the NOSP model. Furthermore, the treatment policy is even more proactive as θ increases.*

In this scenario, the risk associated with the termination decision is eliminated due to the deterministic nature of $G(s)$. However, the system retains inherent risks due to the stochastic nature of health state transitions. As a result, the ROSP model opts for a more assertive stopping policy to mitigate the overall risk.

Next, we explore the case when the optimal policy of ROSP is more conservative. The optimal policy dictates ceasing treatment if $h_t^\theta(s) \le 0$ and continuing otherwise. We say that the policy of ROSP is more conservative than NOSP at time t, if $\mathcal{R}_{t,0}^0 \subset \mathcal{R}_{t,\theta}^0$, where $\mathcal{R}_{t,\theta}^0 = \{s | h_t^\theta(s) > 0\}$. Theorem 3.3 presents a condition such that ROSP consistently adopts more conservative actions.

To establish the proof of the theorem, we first present the following lemma.

Lemma 3.1. *For any $0 \le t \le T-2$, we have*

$$h_t^\theta(s) - h_{t+1}^\theta(s) \ge \sum_{s' \in \mathcal{S}} P(s,s') \left(-U_{t+1}(\theta, s') + U_{t+2}(\theta, s')\right).$$

The following theorem builds on Lemma 3.1, using the recursive structure of the value function.

Theorem 3.3 (Conservative Policy)**.** *If $h_{T-1}^\theta(s) \ge h_{T-1}^0(s)$ for any $s \in \mathcal{R}_{0,0}^0$, i.e.,*

$$\mu_\theta\left(\tilde{G}(s)\right) - \frac{1}{\theta}\log\left(\sum_{s'\in\mathcal{S}} P(s,s')\exp\left(\theta\mu_\theta\left(\tilde{G}(s')\right)\right)\right) \ge \mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s)\right] - \sum_{s'\in\mathcal{S}} P(s,s')\mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s')\right],$$

then the optimal treatment policy for the ROSP will be more conservative than the optimal treatment policy for the NOSP.

This theorem presents a sufficient condition such that the policy for ROSP is always more conservative than NOSP. It is possible that the optimal policy for ROSP spans a spectrum from conservative to aggressive actions, which is in general difficult to analyze. In the following analysis, we consider a simplified case when the policy for NOSP is temporally homogeneous except the final period, *i.e.*, $\mathcal{R}_{0,0}^0 = \mathcal{R}_{1,0}^0, ..., = \mathcal{R}_{T-1,0}^0$.

Corollary 3.1. *If $\mathcal{R}_{0,0}^0 = \mathcal{R}_{T-1,0}^0$, then there are three scenarios of ROSP policies in comparison to NOSP: a more aggressive policy, a more conservative policy, or a policy that transitions from conservative to aggressive.*

This result is a quick application of Theorem 3.1.

3.4 ROSP with predicted states

We expand the ROSP model in this section to integrate predictive information regarding patients' future health conditions. In our enhanced model, DMs, such as clinicians, can utilize both current and predictive information about patients. This allows DMs to be aware not only of the present state s, but also of a forecasted patient class for the next epoch, denoted as $\hat{s}_{t+1} \in \mathcal{S}$. Consequently, our model features an expanded state space, represented as $\mathcal{S} \times \mathcal{S}$. As the state space is extended, the transition probability is defined from state $(s_t, \hat{s}_{t+1})$ to $(s_{t+1}, \hat{s}_{t+2})$. In the field of machine learning, specifically the classification problem, the confusion matrix is widely used to evaluate the performance of predictive models. For example, the DM can use historical classification results to estimate a confusion matrix $\mathbf{Q}$. Each element of the misclassification matrix $Q(i,j) = \mathbb{P}(\hat{s} = j | s = i)$ represents the probability that a patient transitioning to class i at the next epoch is predicted to transition to class j when the prediction is made at the current epoch. The diagonal components of $\mathbf{Q}$ denote the accuracy of the predictive model. If the prediction is perfect, $\mathbf{Q}$ is an identity matrix. Using Bayes' rule, given the current state $(s_t, \hat{s}_{t+1})$, the probability that the actual state at the next epoch is

$$q(s_{t+1}|s_t, \hat{s}_{t+1}) = \frac{P(s_t, s_{t+1})Q(s_{t+1}, \hat{s}_{t+1})}{\tilde{Q}(s_t, \hat{s}_{t+1})},$$

where $\tilde{Q}(s_t, \hat{s}_{t+1}) = \sum_{s_{t+1} \in \mathcal{S}} P(s_t, s_{t+1})Q(s_{t+1}, \hat{s}_{t+1})$, and the vector $\boldsymbol{q}$ can be interpreted as a belief vector of the DM on the next period's state given predicted class $\hat{s}_{t+1}$. Then, the corresponding transition probability is defined as

$$\mathbb{P}(s_{t+1}, \hat{s}_{t+2}|s_t, \hat{s}_{t+1}) = q(s_{t+1}|s_t, \hat{s}_{t+1})\tilde{Q}(s_{t+1}, \hat{s}_{t+2}). \tag{3.9}$$

We use $U_t^{\mathrm{P}}(\theta, s_t, \hat{s}_{t+1})$ to denote the optimal U function of ROSP with predicted future states (ROSP-P). Then the new optimality equation can be expressed as: for any $t \in \mathcal{T} \backslash \{T\}$, $U_t^{\mathrm{P}}(\theta, s_t, \hat{s}_{t+1}) =$

$$\begin{cases} \min\left\{ \mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s_t)\right], H(s_t) + \sum\limits_{s_{t+1}, \hat{s}_{t+2} \in \mathcal{S}} \mathbb{P}(s_{t+1}, \hat{s}_{t+2}|s_t, \hat{s}_{t+1}) \cdot U_{t+1}^{\mathrm{P}}(0, s_{t+1}, \hat{s}_{t+2}) \right\}, & \text{if } \theta = 0, \\ \min\left\{ \mu_\theta(\tilde{G}(s_t)), H(s_t) + \frac{1}{\theta} \log \sum\limits_{s_{t+1}, \hat{s}_{t+2} \in \mathcal{S}} \mathbb{P}(s_{t+1}, \hat{s}_{t+2}|s_t, \hat{s}_{t+1}) e^{\theta \cdot U_{t+1}^{\mathrm{P}}(\theta, s_{t+1}, \hat{s}_{t+2})} \right\}, & \text{else.} \end{cases}$$

The formulation in the final period is $U_T^{\mathrm{P}}(\theta, s_T, \hat{s}_{T+1}) = \mu_\theta\left(\tilde{G}(s_T)\right)$ for all $(s_T, \hat{s}_{T+1}) \in \mathcal{S} \times \mathcal{S}$.

In order to compare ROSP and ROSP-P, we first calculate the expected optimal U function in the ROSP-P model,

$$\bar{U}_t^{\mathrm{P}}(\theta, s_t) = \begin{cases} \sum_{\hat{s}_{t+1} \in \mathcal{S}} \tilde{Q}(s_t, \hat{s}_{t+1}) U_t^{\mathrm{P}}(0, s_t, \hat{s}_{t+1}) & \theta = 0, \\ \frac{1}{\theta} \log \sum_{\hat{s}_{t+1} \in \mathcal{S}} \tilde{Q}(s_t, \hat{s}_{t+1}) \exp\left(\theta \cdot U_t^{\mathrm{P}}(\theta, s_t, \hat{s}_{t+1})\right) & \theta > 0. \end{cases} \tag{3.10}$$

If we have an initial distribution $\boldsymbol{p}_0$, the expected outcome can be calculated via

$$\frac{1}{\theta} \log \sum_{s \in \mathcal{S}} p_0(s) \exp\left(\theta \cdot \bar{U}_0^P(\theta, s)\right)$$

Then, we can derive the following result.

Proposition 3.6. *The optimal policy of ROSP-P consistently outperforms the one of ROSP, i.e., $\bar{U}_t^{\mathrm{P}}(\theta, s_t) \leq U_t(\theta, s_t)$, for any $\theta \in [0, +\infty)$, $t \in \mathcal{T}$ and $s_t \in \mathcal{S}$.*

Proposition 3.6 suggests that any predictive information (even if inaccurate) is valuable for our ROSP model. That is, for any patient state s, the risk measure of cost under the optimal policy of the model with predictive information is not higher than that under the optimal policy of the model without predictive information. This result is based on the fact that the action can be taken on the extended state space with the predicted state, and the true misclassification matrix $\mathbf{Q}$ is known to the DM.

We next proceed to analyze the structure of the utility function in our ROSP model with predictive information. To establish the desired results, we introduce the following assumption regarding the misclassification matrix.

Assumption 3.4. *The misclassification matrix $\mathbf{Q}$ is TP2.*

Under this assumption, and given that the transition matrix $\mathbf{P}$ is also TP2, the product matrix $\tilde{\mathbf{Q}} = \mathbf{P} \cdot \mathbf{Q}$ inherits the TP2 property. This preservation of the TP2 property under matrix multiplication is a well-established result in the literature (see, for example, Lemma 10.5.2 in Krishnamurthy (2016)). Leveraging this property, we derive the structure of the utility function $U_t^{\mathrm{P}}(\theta, s_t, \hat{s}_{t+1})$.

Proposition 3.7. *Under Assumptions 3.1, 3.2 and 3.4, the value function $U_t^{\mathrm{P}}(\theta, s_t, \hat{s}_{t+1})$ is decreasing in s_t and $\hat{s}_{t+1}$ for any $t \in \mathcal{T}$ and $\theta \in \mathbb{R}_+$.*

In the context of ROSP-P, we can define the benefit function of continuing treatment as follows:

$$h_t^{\mathrm{P},\theta}(s,\hat{s}) = \begin{cases} \mu_\theta\left(\tilde{G}(s)\right) - H(s) - \dfrac{1}{\theta}\log\displaystyle\sum_{s',\hat{s}'\in\mathcal{S}} \mathbb{P}(s',\hat{s}'|s,\hat{s})\exp\left(\theta\cdot U_{t+1}^{\mathrm{P}}(\theta,s',\hat{s}')\right) & \text{if } \theta>0, \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s)\right] - H(s) - \displaystyle\sum_{s',\hat{s}'\in\mathcal{S}} \mathbb{P}(s',\hat{s}'|s,\hat{s})U_{t+1}^{\mathrm{P}}(\theta,s',\hat{s}') & \text{if } \theta=0. \end{cases} \tag{3.11}$$

Similarly, we stop the treatment if $h_t^{\mathrm{P},\theta}(s,\hat{s}) \leq 0$. Based on Proposition 3.7, we can elucidate the characteristics of the optimal policy's structure.

Proposition 3.8. *Under Assumptions 3.1, 3.2 and 3.4, we have the following two statements:*

1. *The benefit $h_t^{P,\theta}(s,\hat{s})$ of continuing the treatment is increasing in $\hat{s}$.*

2. *At each epoch $t\in\mathcal{T}$, there exists an optimal switching curve $s_t \mapsto \tau_t^\theta(s_t)$ that partitions the state space $\mathcal{S}^2$ into two subsets w.r.t. termination and continuation, i.e.,*

$$a_t^\star(s,\hat{s}) = \begin{cases} 1 & \hat{s}\leq \tau_t^\theta(s), \\ 0 & \hat{s} > \tau_t^\theta(s). \end{cases}$$

Proposition 3.8 establishes the threshold structure within the optimal policy of ROSP-P. Such kind of switching curves are visually illustrated in Figure 3.4.

3.5 Case studies

In this section, we investigate the performance of our algorithms, ROSP and ROSP-P, in different settings.

3.5.1 Extubation decision

Within critical care context, the decision-making process for extubation of mechanically ventilated patients, who are unable to breathe independently, emerges as a critical challenge. Prolonged intubation in ICUs incurs substantial costs, impacting both patient well-being and hospital resources (Barrett, et al., 2014; Halpern and Pastores, 2015). Moreover, the timing of extubation is often regarded as a crucial metric for evaluating the quality of surgical care in hospitals (Wong, et al., 2016).

Early extubation, while potentially reducing ventilator time for patients, carries inherent risks of failure. Such failures may result in the need for re-intubation or the employment of alternative respiratory support mechanisms, particularly in instances where spontaneous respiration is not achieved, which is correlated with prolonged ICU durations and an elevation in mortality rates (Epstein, et al., 1997; Thille, et al., 2013). On the other hand, extended periods of mechanical ventilation can overburden ICU resources and exacerbate unit congestion. The challenges of resource allocation, particularly concerning ventilators, were starkly highlighted during the COVID-19 pandemic, as physicians were faced with arduous decisions regarding ventilator allocation among critically ill patients (Cohen, et al., 2020).

To optimize extubation decisions dynamically upon the evolving health state of patients, it is natural to apply a MDP which has seen widespread use in the healthcare domain when dealing with dynamic and stochastic decision-making problems (Denton, et al., 2009; Chhatwal, et al., 2010; Ayvaci, et al., 2012). The standard optimization criterion typically involves minimizing the expected cost, which means finding a policy that minimizes the value function given the system's initial state. However, in many applications, there's a preference to also minimize a measure of risk alongside this standard criterion. This requires a criterion that accounts for the variability induced by a given policy, due to the stochastic nature of the system.

3.5.1.1 Extubation without prediction.

Firstly, we consider a five-class (state) model (*i.e.*, $\mathcal{S} = \{1, 2, 3, 4, 5\}$) for extubation and use the parameters estimated in Cheng, et al. (2024). The immediate continuation cost, denoted as H, corresponds to the decision interval, set at a duration of 6 hours. This specific time frame is chosen based on the practice, where medical staff assesses the condition of patients on ventilators at 6-hour intervals to determine the appropriateness of extubation.

The transition probability is estimated as follows:

$$\hat{\boldsymbol{P}} = \begin{pmatrix} 0.35 & 0.21 & 0.20 & 0.11 & 0.13 \\ 0.09 & 0.40 & 0.28 & 0.13 & 0.10 \\ 0.02 & 0.14 & 0.48 & 0.24 & 0.12 \\ 0.01 & 0.05 & 0.23 & 0.36 & 0.35 \\ 0.01 & 0.02 & 0.07 & 0.18 & 0.72 \end{pmatrix}. \tag{3.12}$$

Terminal Cost. The terminal cost $\tilde{G}(s)$ is determined based on the Remaining length of stay (RLOS) following extubation. As demonstrated by Cheng, et al. (2024), the expected RLOS for five distinct patient classes was estimated, with the results summarized in Table 3.1. Notably, the RLOS exhibits a monotonic decrease as the state s progresses, aligning with clinical expectations.

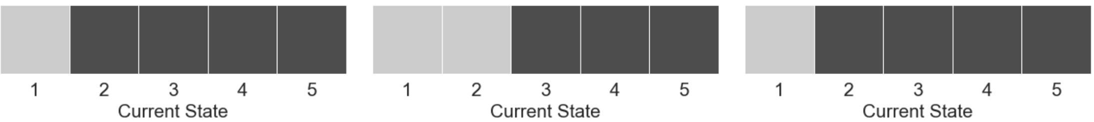
To further assess the performance of the model, additional distributional information is required. For this purpose, synthetic data generation methods are employed. Specifically, the

Table 3.1: Expected RLOS Conditioned on Patient Class

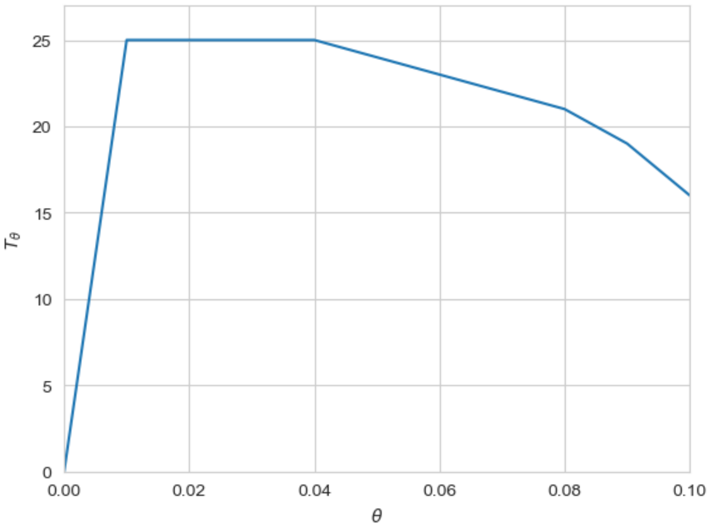
Terminal Class	1	2	3	4	5
RLOS (hours)	137.8	98.2	91.1	88.6	79.0

RLOS samples are assumed to follow a Gaussian distribution, truncated to ensure non-negativity. The standard deviations for the respective classes are set to 60, 60, 45, 40, and 30, reflecting a decreasing trend in line with Assumption 3.1.

Comparison to NOSP. Then, we examine how our approach differs from the traditional nominal optimal stopping problem. Unlike NOSP, which aims to minimize the expected total cost, our ROSP framework focuses on the risk measure parameterized risk tolerance θ. For analysis, we assume an initial distribution of $p_0 = (0.5, 0.5, 0, 0, 0)$ and select θ from $\{0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1\}$. The optimal policies for both NOSP and ROSP across various targets are depicted in Figure 3.2. According to NOSP, patients are extubated as soon as they improve to class 2 or higher, which is consistent with findings from McConville and Kress (2012). In contrast, ROSP, which accounts for risk considerations, tends towards a more conservative stance, opting for extubation in class 3 or above during the initial time frames (from the initial period to period T_θ). Notably, as θ increases from small values, the policy becomes increasingly conservative at the outset, reflecting a cautious approach to early extubation. However, beyond a certain threshold, further increases in θ leads to a more aggressive policy, noticed in Figure 3.2 (d).



(a) NOSP $t = 0, ..., 26$ **(b)** ROSP $t = 0, ..., T_\theta$ **(c)** ROSP $t = T_\theta, ..., 26$



(d) ROSP: T_θ w.r.t. different θ

Figure 3.2: Optimal policies of NOSP and ROSP. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.

To assess the effectiveness of the proposed policies, we conducted Monte Carlo simulations, generating $N = 10^6$ sample trajectories for each policy. This approach enabled the calculation of the total cost for each trajectory, resulting in 10^6 realized costs per policy, denoted by $\{{}^j V\}$. We evaluated performance using two metrics: α-VaR and α-CVaR with α selected from $\{0.2, 0.1, 0.05, 0.01\}$. The outcomes are detailed in Table 3.2, where each column reflects outcomes for different α values. We observed that varying the parameter θ primarily influences decisions in the later periods, which has a negligible impact on the final outcomes, as the results remain consistent with two decimal places. Consequently, we report the result for ROSP with $\theta = 0.02$.

Table 3.2: Risk profile of NOSP and ROSP

	$\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{V}]$	$\mathrm{VaR}_{0.2}$	$\mathrm{VaR}_{0.1}$	$\mathrm{VaR}_{0.05}$	$\mathrm{VaR}_{0.01}$	$\mathrm{CVaR}_{0.2}$	$\mathrm{CVaR}_{0.1}$	$\mathrm{CVaR}_{0.05}$	$\mathrm{CVaR}_{0.01}$
NOSP	98.99	143.01	168.31	190.00	231.80	175.73	196.88	215.61	252.81
ROSP	100.13	134.51	154.04	170.73	202.68	159.71	175.97	190.32	218.81

The expected costs $\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{V}\right]$ under NOSP and ROSP were 98.99 and 100.13, respectively. As illustrated in the results, ROSP incurs a slight increase in expected cost (approximately 1.15%) compared to the NOSP framework. However, the modest cost increase is accompanied by significant improvements in risk profiles. Specifically, the ROSP framework achieves a 12.56% reduction in $\mathrm{VaR}_{0.01}$ and a 13.45% reduction in $\mathrm{CVaR}_{0.01}$. Furthermore, the performance of the ROSP framework becomes more pronounced as threshold α increases. This highlights that ROSP focuses more on controlling extreme values, making it effective in managing tail risks.

3.5.1.2 Extubation with predictive information.

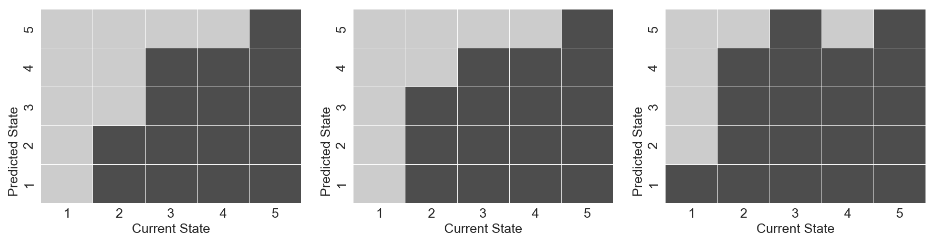
To capture the state of prediction and associated errors, we utilize a misclassification matrix, symbolized as $\boldsymbol{Q}$. This matrix serves as a counterpart to the confusion matrix in multi-class classification problems. Each element within the misclassification matrix $\boldsymbol{Q}(i,j) := \mathbb{P}(\hat{s}_t = j | s_t = i)$ signifies the likelihood of a patient from class i being erroneously predicted as class j. The probabilities of correctly predicting each class of patients are reflected in the diagonal elements of $\boldsymbol{Q}$, akin to the sensitivity and specificity measures in a binary classification matrix. In an ideal scenario of perfect prediction accuracy, $\boldsymbol{Q}$ would be an identity matrix. Presented below is the out-of-sample misclassification matrix $\hat{\boldsymbol{Q}}$ as derived from a predictive model:

$$\hat{\boldsymbol{Q}} = \begin{pmatrix} 0.60 & 0.08 & 0.17 & 0.15 & 0.00 \\ 0.10 & 0.60 & 0.16 & 0.09 & 0.05 \\ 0.01 & 0.09 & 0.61 & 0.21 & 0.08 \\ 0.00 & 0.02 & 0.31 & 0.53 & 0.14 \\ 0.00 & 0.01 & 0.11 & 0.28 & 0.60 \end{pmatrix}. \tag{3.13}$$

For ROSP-P, we select θ from the set $\{0.005, 0.01, 0.015, 0.02, 0.025, 0.3\}$. The optimal policies for NOSP with predicted future states (NOSP-P) and ROSP-P with $\theta = 0.02$ are illustrated in Figures 3.3 and 3.4, respectively, while the optimal policies under other values of θ are provided in Appendix A.6. These figures depict the optimal policy at each decision period, contingent on both the current patient class and predicted future classes. The incorporation of prediction information enhances the precision of the extubation strategy by extending the state space.

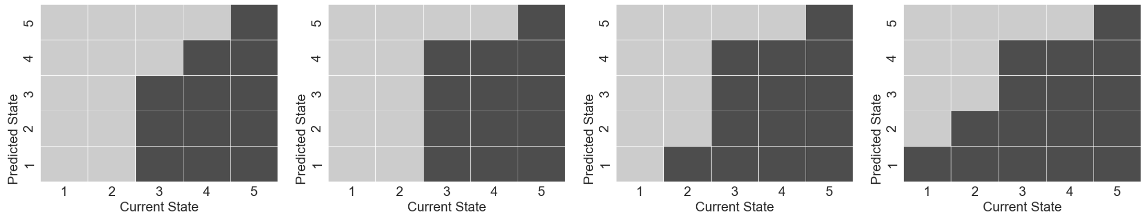
A notable observation is that decision-making becomes more aggressive as time t progresses, a phenomenon driven by the terminal effect. The results for NOSP-P, as shown in Figure 3.3, align with those reported in Cheng, et al. (2024). In contrast, ROSP-P, which incorporates risk information, exhibits a more conservative policy compared to NOSP-P, consistent with the findings in Section 3.5.1.1. Specifically, ROSP-P adopts a more cautious approach in the initial periods to assess treatment efficacy. For instance, decision-makers retain patients in classes (3, 4) for an extended duration when $t \leq 17$, and similarly, patients in classes (2, 1) are retained longer when $t \leq 23$.

Varying the parameter θ reveals different levels of conservativeness, with policies becoming increasingly conservative as θ increases. For example, when $\theta = 0.005$, ROSP-P recommends a continuation strategy for class 2 patients predicted to transition to class 3 at epochs 24 and 25. Additionally, it suggests extending treatment for one additional period for patients in states $(2, 3)$, $(2, 4)$, and $(3, 5)$ at epoch 26. Compared to NOSP-P, this results in ROSP-P altering its policy actions at a total of 5 nodes within the tuple $(s_t, \hat{s}_{t+1}, t)$. As θ increases to 0.015, policy modifications occur at 53 nodes. With $\theta = 0.02$, ROSP-P opts to continue treatment for all class 2 patients through epochs 0 to 23 and for class 3 patients predicted to progress to class 4 during epochs 0 to 17, in contrast to the policy of NOSP-P. At epochs 24 and 25, continuation is recommended for class 2 patients predicted to transition to classes 2 or 3, while at epoch 26, an additional period of therapy is advocated for patients in states $(2, 3)$, $(2, 4)$, and $(3, 5)$. Overall, this leads to a total of 55 nodes at which ROSP-P diverges from the NOSP-P policy. Further increasing θ to 0.025 and 0.03 results in policy modifications at 73 and 78 nodes, respectively, demonstrating a clear trend toward greater conservativeness with higher values of θ.



(a) $t = 0, .., 23$ **(b)** $t = 24, 25$ **(c)** $t = 26$

Figure 3.3: Optimal policy of NOSP-P. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.



(a) $t = 0, ..., 17$ **(b)** $t = 18, ..., 23$ **(c)** $t = 24, 25$ **(d)** $t = 26$

Figure 3.4: Optimal policy of ROSP-P given $\theta = 0.02$. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.

Lastly, we utilize Monte Carlo simulations to generate 10^6 samples to evaluate the performance of NOSP-P and ROSP-P. In this numerical case study, distinct policies emerge for different target values, attributable to the refined extubation strategy within the expanded state space. The expected total costs under varying parameters θ are $98.09, 100.07,$ and 100.70. Note that ROSP-P sacrifices greater expected outcome for improved risk profile as θ increases. This is paralleled by an increasingly conservative extubation strategy. Then we compute VaR and CVaR in Table 3.3.

Table 3.3: Risk profile of NOSP-P and ROSP-P

	$\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{V}]$	$\mathrm{VaR}_{0.2}$	$\mathrm{VaR}_{0.1}$	$\mathrm{VaR}_{0.05}$	$\mathrm{VaR}_{0.01}$	$\mathrm{CVaR}_{0.2}$	$\mathrm{CVaR}_{0.1}$	$\mathrm{CVaR}_{0.05}$	$\mathrm{CVaR}_{0.01}$
NOSP-P	98.09	134.87	157.02	176.67	216.59	164.25	183.56	201.23	238.01
ROSP-P									
$\theta = 0.005$	98.09	134.91	156.98	176.60	216.54	164.22	183.51	201.18	238.06
$\theta = 0.01$	100.07	133.31	152.46	168.97	201.08	158.18	174.30	188.67	217.50
$\theta = 0.015$	100.07	133.31	152.46	168.97	201.08	158.18	174.30	188.67	217.50
$\theta = 0.02$	100.70	133.26	152.02	168.28	200.04	157.71	173.58	187.80	216.53
$\theta = 0.025$	100.70	133.26	152.02	168.28	200.04	157.71	173.58	187.80	216.53
$\theta = 0.03$	100.70	133.26	152.02	168.28	200.04	157.71	173.58	187.80	216.53

The empirical observations align coherently with the analytical findings presented in Proposition 3.6. When benchmarked against their respective standard NOSP and ROSP frameworks, both NOSP-P and ROSP-P demonstrate superior performance. This enhanced efficacy is evident across key metrics, including violation probabilities and conditional expected shortages, and can be attributed to the incorporation of predictive information.

A comparative analysis between NOSP-P and ROSP-P reveals that as the risk parameter θ increases, ROSP-P achieves a more favorable risk profile, albeit with a marginal increase in expected costs. Specifically, when comparing NOSP-P and ROSP-P with $\theta = 0.005$, the difference in performance is negligible. However, as θ is increased to 0.01, a significant improvement in risk management is observed. Further increases in θ intensify the magnitude of this improvement. Notably, the policy undergoes major modifications in the later periods, resulting in negligible changes to the distributions of total costs. These findings reaffirm the previously observed trade-off between expected performance and risk profile, consistent with the behavior observed in models without predictive information.

3.5.2 Discharge decision with readmissions

To further evaluate the performance of the ROSP model, we conducted experiments using the MIMIC-IV database (Johnson, et al., 2020). Our study focuses on the discharge decision in the Medical Intensive Care Unit (MICU) and studies the trade-off between current workload and readmission risk. Detailed descriptions of clinical and demographic variables we use from MIMIC-IV , as well as parameter estimation process, are provided in Appendix A.5. We model the sequential organ failure assessment (SOFA) score as the state variable with an additional state to indicate the in-ICU death.

Problem settings. In our study, physicians assess a patient's condition every 12 hours to determine discharge decisions, which may involve transfers to the general ward (or step-down units). The maximum duration T is set as 27 (13.5 days), while long-stayers are not the focus of this study. In our dataset, only 4.70% of patients has LOS in MICU exceeding 13.5 days. Due to the limited availability of data for high SOFA scores (see Figure A.5 in Appendix A.5), we aggregate all observations with SOFA scores ≥ 16 into a single state. The terminal cost is modeled as the weighted total cost of mortality and readmission workload costs, formulated as follows:

$$\tilde{G}(s) = c_{mort} \cdot \mathrm{B}in(1, p_{mort}(s)) + c_{re} \cdot \mathrm{B}in(1, p_{re}(s)) \cdot Log\text{–}\mathcal{N}(\mu_r, \sigma_r^2),$$

where $p_{mort}(s)$ and $p_{re}(s)$ represent the 30-day mortality rate and readmission rate, respectively, for patients discharged in state s. We set $c_{mort} = 50$ and $c_{re} = 10$ for this numerical study, while the cost of the first admission is two unit costs per day, *i.e.*, $H(s) = 1$. Using readmitted patient data from MIMIC-IV , we estimate the parameters for readmitted LOS as $\mu_r = 0.87$ and $\sigma_r = 1.05$. We further truncate the distribution on $[0.1, 13.95]$, where 13.95 is the 95th quantile of readmitted LOS distribution. This truncated Log-normal distribution ensures tractability. The remaining parameters are estimated using MIMIC-IV data and the procedure is discussed in Appendix A.5. The initial state distribution assumes that patients are uniformly distributed across states 10 to 16+.

Table 3.4: States for which discharge decision is optimal in NOSP and ROSP

t	0	1-9	10-13	14-17	18 and 19	20	21-23	24	25	26
NOSP	0-9	0-9	0-9	0-9	0-10	0-10	0-10	0-10	0-11	0-11, 15
ROSP, $\theta = 0.05$	0-6	0-6	0-7	0-7	0-7	0-8	0-8	0-8	0-9	0-9, 15
ROSP, $\theta = 0.1$	0-7	0-8	0-8	0-9	0-9	0-9	0-10	0-11	0-11	0-12, 14, 15

Results. We compare the results between NOSP and ROSP to explore the insights into the effect of risk on hospital discharge decisions. ROSP conducts more conservative policies, particularly in the early decision periods. Compared to policies with $\theta = 0.05$, policies with $\theta = 0.1$ exhibit more aggressive policies, especially in later periods. Using Monte-Carlo simulation with

Table 3.5: Risk profile of total cost of discharge problem

	$\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{V}]$	$\text{VaR}_{0.1}$	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{CVaR}_{0.1}$	$\text{CVaR}_{0.05}$	$\text{CVaR}_{0.01}$
NOSP	31.33	63	81.05	92.29	73.53	95.21	104.34
ROSP, $\theta = 0.05$	31.76	64	76	84.87	71.89	88.49	98.82
ROSP, $\theta = 0.1$	32.06	64	76	83.42	71.56	87.53	97.70

10^6 samples, we evaluate the risk profile, including VaR and CVaR, as summarized in Table 3.5. The results demonstrate that the ROSP model provides superior control over tail distributions, effectively mitigating the risk of high-cost outcomes, which typically correspond to patient mortality. As the risk sensitivity parameter θ increases, the ROSP model achieves better risk control at the expense of slightly degraded expected performance.

3.6 Conclusion

Our research studies a novel risk measure-based framework within a finite-horizon discrete-time Markov chain, specifically designed to address medical optimal stopping problems through the lens of risk trade-offs. Central to our methodology is the incorporation of an additional dimension: the risk tolerance level (denoted by θ), which is absent in nominal MDP approaches that overlook risk profiles. We present two models, ROSP and ROSP-P, where ROSP-P incorporates the predictive information.

To investigate the impact on risk profiles, we analyze the policy structure of our risk-aware models. We showcase the submodularity of the utility function and the threshold structure of the optimal policies under some mild conditions. Subsequently, we compare the policies between NOSP and ROSP, demonstrating the conditions under which our models suggest more aggressive or conservative policies. We identify three possible scenarios of optimal strategies in a special case. The findings reveal a nuanced trade-off between stopping and continuing risks in healthcare decision-making.

To validate the proposed framework, we apply our models, ROSP and ROSP-P, to two stopping problems in ICUs. Through extensive numerical experiments across various settings, the results demonstrate significant improvements in risk profiles, such as VaR and CVaR, compared to NOSP (and NOSP-P, respectively). Consequently, the proposed policies achieve superior risk profiles with only a marginal loss in expected performance.

Our research can be extended in several directions. First, while our analysis of structural properties focuses on a one-dimensional state space, future work could explore higher-dimensional state spaces, where the state space may exhibit a partial order with respect to cost functions rather than a total order. Second, this paper examines risk-sensitive MDPs in the context of optimal stopping for medical treatment. In practice, numerous decisions arise during therapy procedures. Investigating risk-sensitive MDPs across a broader spectrum of discrete decision-making scenarios in therapy could yield valuable insights.

In conclusion, this study advances the field of healthcare analytics by providing a deeper understanding of risk management in patient care decisions. It opens avenues for future research to explore other aspects of healthcare management using similar decision process frameworks, with potential applications extending to public health and medical resource allocation.

Appendix

A.1 Proofs

Proof of Proposition 3.2

Proof. Consider an arbitrary policy $\boldsymbol{\pi} = (\pi_0, \cdots, \pi_{T-1})$ where π_t maps s_t to action a_t for $t = 0, \cdots, T-1$.

Accordingly, the risk-to-go of an arbitrary policy $\boldsymbol{\pi}$ can be expressed recursively as

$$U_t^{\boldsymbol{\pi}}(\theta, s) = \begin{cases} \mu_\theta\left(\tilde{G}(s)\right) & \text{if } \pi_t(\theta, s) = 1, \\ H(s) + \dfrac{1}{\theta} \log \displaystyle\sum_{s'} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}^{\boldsymbol{\pi}}(\theta, s')\right) & \text{if } \pi_t(\theta, s) = 0. \end{cases}$$

The principle of optimality is equivalent to the following two **statements**:

1. The deterministic Markovian policy $\boldsymbol{\pi}^*$ obtained via the stochastic dynamic programming recursion Equation (Eq.)(3.6) is optimal in the sense of Eq.(3.4).

2. Conversely, the optimal risk-to-go, *i.e.*, $\min_{\boldsymbol{\pi}} U_t^{\boldsymbol{\pi}}(\theta, s)$, satisfies the optimal dynamic programming recursion Eq.(3.6).

Then, we prove these two statements.

Statement 1: To establish statement 1, we show by backward induction for $t = T, \ldots, 0$ that $U_t^{\boldsymbol{\pi}}(\theta, s) \geq U_t(\theta, s)$ with equality achieved at $\boldsymbol{\pi} = \boldsymbol{\pi}^*$ where $\boldsymbol{\pi}^*$ is obtained via Eq.(3.6).

When $t = T$, $\mu_\theta\left(\tilde{V}_T^{\boldsymbol{\pi}}(s)\right) = \mu_\theta\left(\tilde{G}(s)\right) = U_T(\theta, s)$. Assume that $\mu_\theta(\tilde{V}_{t+1}^{\boldsymbol{\pi}}(s)) \geq U_{t+1}(s),\ \forall s \in \mathcal{S}$, for some $t+1$, then

$$
\begin{aligned}
\mu_\theta\left(\tilde{V}_t^{\boldsymbol{\pi}}(s)\right) &= \begin{cases} \mu_\theta\left(\tilde{G}(s)\right) & \text{if } \pi_t(\theta, s) = 1, \\ H(s) + \dfrac{1}{\theta}\log\sum_{s'} P(s, s')\cdot \exp\left(\theta\cdot\mu_\theta\left(\tilde{V}_{t+1}^{\boldsymbol{\pi}}(s')\right)\right) & \text{if } \pi_t(\theta, s) = 0. \end{cases} \\
&\geq \begin{cases} \mu_\theta\left(\tilde{G}(s)\right) & \text{if } \pi_t(\theta, s) = 1, \\ H(s) + \dfrac{1}{\theta}\log\sum_{s'} P(s, s')\cdot \exp\left(\theta\cdot U_{t+1}(\theta, s')\right) & \text{if } \pi_t(\theta, s) = 0. \end{cases} \\
&\geq \min\left\{\mu_\theta(\tilde{G}(s)), H(s) + \frac{1}{\theta}\log\sum_{s_{t+1}} P(s, s_{t+1})\cdot\exp\left(\theta\cdot U_{t+1}(\theta, s_{t+1})\right)\right\} \\
&= U_t(\theta, s).
\end{aligned}
$$

By mathematical induction, we prove that $\mu_\theta\left(\tilde{V}_0^{\boldsymbol{\pi}}(s)\right) \geq U_0(\theta, s)$, which implies $\min_{\boldsymbol{\pi}} \mu_\theta\left(\tilde{V}_0^{\boldsymbol{\pi}}(s)\right) \geq U_0(\theta, s)$. The equality holds when $\boldsymbol{\pi} = \boldsymbol{\pi}^*$. In this regard, $\boldsymbol{\pi}^*$ is the optimal solution of Eq. (3.4).

Statement 2: Define $U_t^*(\theta, s) = \min_{\boldsymbol{\pi}} U_t^{\boldsymbol{\pi}}(\theta, s)$. We need to show that $U_t^*(\theta, s)$ satisfies Bellman's equation, *i.e.*,

$$
U_t^*(\theta, s) = \min\left\{\mu_\theta\left(\tilde{G}(s)\right), H(s) + \frac{1}{\theta}\log\sum_{s'} P(s, s')\cdot\exp\left(\theta\cdot U_{t+1}^*(\theta, s')\right)\right\}. \tag{A.14}
$$

To prove statement 2, we first show that the Left hand side (LHS) of Eq. (A.14) is greater than the Right hand side (RHS).

$$
\begin{aligned}
U_t^*(\theta, s) &= \min_{\boldsymbol{\pi}}\left\{\mu_\theta\left(\tilde{V}_t^{\boldsymbol{\pi}}(s)\right)\right\} \\
&= \min\left\{\mu_\theta(\tilde{G}(s)), \min_{\boldsymbol{\pi}}\left\{H(s) + \frac{1}{\theta}\log\sum_{s'} P(s, s')\cdot\exp\left(\theta\cdot\mu_\theta\left(\tilde{V}_{t+1}^{\boldsymbol{\pi}}(s')\right)\right)\right\}\right\} \\
&\geq \min\left\{\mu_\theta\left(\tilde{G}(s)\right), H(s) + \frac{1}{\theta}\log\sum_{s'} P(s, s')\cdot\exp\left(\theta\cdot\min_{\boldsymbol{\pi}_{s'}}\mu_\theta\left(\tilde{V}_{t+1}^{\boldsymbol{\pi}_{s'}}(s')\right)\right)\right\} \\
&= \min\left\{\mu_\theta\left(\tilde{G}(s)\right), H(s) + \frac{1}{\theta}\log\sum_{s'} P(s, s')\cdot\exp\left(\theta\cdot U_{t+1}^*(\theta, s)\right)\right\},
\end{aligned} \tag{A.15}
$$

On the other hand, we need to prove the LHS is greater than the LHS. Suppose $\boldsymbol{\pi}_{t:}^* = (\pi_t^*, \pi_{t+1}^*, \ldots, \pi_{T-1}^*)$ is the optimal solution of U_t^*. Let $\hat{\boldsymbol{\pi}}_{t:} = (\pi_t, \pi_{t+1}^*, \ldots, \pi_{T-1}^*)$, where the policy π_t is arbitrary. According to the definition of $\boldsymbol{\pi}^*$,

$$
\mu_\theta\left(\tilde{V}_t^{\boldsymbol{\pi}_{t:}^*}(s)\right) \leq \mu_\theta\left(\tilde{V}_t^{\hat{\boldsymbol{\pi}}_{t:}}(s)\right) \tag{A.16}
$$

As π_t is arbitrary, we take a minimization on the LHS.

$$U_t^*(\theta, s) \leq \min\left\{\mu_\theta\left(\tilde{G}(s)\right), H(s) + \frac{1}{\theta}\log\sum_{s'} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}^*(\theta, s)\right)\right\} \tag{A.17}$$

Thus, we complete the proof of Proposition 3.2.

□

Proof of Proposition 3.3

Proof. (a) First, we use the properties of Entropic risk measure (Appendix A.2) to show that $U_t(\theta, s)$ is increasing in θ. For any $\theta_1 \leq \theta_2$ and $t \in \mathcal{T},\ s \in \mathcal{S}$,

$$U_t(\theta_2, s) = \mu_{\theta_2}\left(\tilde{V}_t^{\boldsymbol{\pi}(\theta_2)}(s)\right) \geq \mu_{\theta_1}\left(\tilde{V}_t^{\boldsymbol{\pi}(\theta_2)}(s)\right) \geq \min_{\boldsymbol{\pi}} \mu_{\theta_1}\left(\tilde{V}_t^{\boldsymbol{\pi}}(s)\right) = U_t(\theta_1, s),$$

where the first inequality is due to the monotonicity in θ and the second inequality is due to the definition of optimal risk-to-go. Then, we use the induction to prove the monotonicity in t. When $t = T - 1$,

$$U_{T-1}(\theta, s) = \min\left\{\mu_\theta\left(\tilde{G}(s)\right), H(s) + \frac{1}{\theta}log\left(\sum_{s' \in \mathcal{S}} P(s, s') \exp\left(\theta \cdot U_T(\theta, s')\right)\right)\right\} \leq U_T(\theta, s).$$

Assume that $U_t(\theta, s) \leq U_{t+1}(\theta, s),\ \forall s \in \mathcal{S}$. Considering the decision period $t-1$, we could have

$$\begin{aligned}
&U_{t-1}(\theta, s) \\
&= \min\left\{\mu_\theta(\tilde{G}(s)),\ H(s) + \frac{1}{\theta} \cdot \log\sum_{s' \in \mathcal{S}} P(s, s') \cdot \exp\left(\theta \cdot U_t(\theta, s')\right)\right\} \\
&\leq \min\left\{\mu_\theta(\tilde{G}(s)),\ H(s) + \frac{1}{\theta} \cdot \log\sum_{s' \in \mathcal{S}} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}(\theta, s')\right)\right\} \\
&= U_t(\theta, s)
\end{aligned}$$

By mathematical induction, we prove the monotonicity in t.

(b) When $t = T$, $U_T(\theta, s) = \mu_\theta\left(\tilde{G}(s_T)\right)$, which is decreasing in s_T according to Assumption 3.1. Then, we use the induction to prove part (b). Assume that in the time epoch $t + 1$, $U_{t+1}(\theta, s_{t+1})$ is decreasing in s_{t+1}.

Then, in the decision epoch t, we could have

$$U_t(\theta, s_t) = \min\left\{\mu_\theta\left(\tilde{G}(s_t)\right),\ H(s_t) + \frac{1}{\theta} \cdot \log\sum_{s_{t+1} \in \mathcal{S}} P(s_t, s_{t+1}) \cdot \exp\left(\theta \cdot U_{t+1}(\theta, s_{t+1})\right)\right\}.$$

Then due to the first-order stochastic dominance and the decreasing assumption of $\exp\left(\theta \cdot (U_{t+1}(\theta, s'))\right)$, we have $\log \sum_{s_{t+1}} P(s_t, s_{t+1}) \cdot \exp\left(\theta \cdot U_{t+1}(\theta, s_{t+1})\right)$ is decreasing in s_t. Since the first term $\mu_\theta\left(\tilde{G}(s_t)\right)$ and the element $H(s_t)$ are also decreasing in s_t, we get $U_t(\theta, s_t)$ is decreasing in s_t, which completes the proof. □

Proof of Proposition 3.4

Proof. Proposition 3.3 shows that $U_t(\theta, s) \leq U_{t+1}(\theta, s)$, for any $s \in \mathcal{S},\ t \in \mathcal{T} \backslash \{T\}$. Based on this, we can have that for any $t \in \mathcal{T}\backslash\{T\}$,

$$\begin{aligned}
h_t^\theta(s) &= \mu_\theta\left(\tilde{G}(s)\right) - H(s) - \frac{1}{\theta} \cdot \log \sum_{s'} P(s, s') \cdot \exp\left(\theta \cdot U_{t+1}(\theta, s')\right) \\
&\geq \mu_\theta\left(\tilde{G}(s)\right) - H(s) - \frac{1}{\theta} \cdot \log \sum_{s'} P(s, s') \cdot \exp\left(\theta \cdot U_{t+2}(\theta, s')\right) \\
&= h_{t+1}^\theta(s).
\end{aligned}$$

Therefore, we complete the proof for proposition 3.4. □

Proof of Proposition 3.5

Proof. This is a quick result based on Theorem 1.C.20. from Shaked and Shanthikumar (2007): Let $\tilde{x}$ and $\tilde{y}$ be independent random variables. Then $\tilde{x} \leq_{lr} \tilde{y}$ if, and only if,

$$\phi(\tilde{x}, \tilde{y}) \leq_{st} \phi(\tilde{y}, \tilde{x})$$

for any $\phi \in \mathcal{G}_{lr} := \{\phi : \mathbb{R}^2 \to \mathbb{R} : \phi(x, y) \leq \phi(y, x) \text{ whenever } x \leq y\}$.

Notice Eq. (3.8) is equivalent to

$$\begin{aligned}
&\frac{\sum_{s' \in \mathcal{S}} P(i, s') \exp\left(\theta f_{t+1}(s')\right)}{\sum_{s' \in \mathcal{S}} P(j, s') \exp\left(\theta f_{t+1}(s')\right)} \geq \frac{\sum_{s' \in \mathcal{S}} P(i, s') \exp\left(\theta f_t(s')\right)}{\sum_{s' \in \mathcal{S}} P(j, s') \exp\left(\theta f_t(s')\right)} \\
\Longleftrightarrow &\mathbb{E}_\mathbb{P}\left[\exp\left(\theta f_{t+1}(\tilde{z}_j)\right)\right] \cdot \mathbb{E}_\mathbb{P}\left[\exp\left(\theta f_t(\tilde{z}_i)\right)\right] \leq \mathbb{E}_\mathbb{P}\left[\exp\left(\theta f_{t+1}(\tilde{z}_i)\right)\right] \cdot \mathbb{E}_\mathbb{P}\left[\exp\left(\theta f_t(\tilde{z}_j)\right)\right].
\end{aligned}$$

We define $\tilde{x} \overset{s.t.}{=} \tilde{z}_i$ and $\tilde{y} \overset{s.t.}{=} \tilde{z}_j$, while $\tilde{x}$ and $\tilde{y}$ are independent. Define $\phi(\tilde{x}, \tilde{y}) = \exp\left(\theta f_t(\tilde{x}) + \theta f_{t+1}(\tilde{y})\right)$, where

$$\begin{aligned}
&\phi(x, y) \leq \phi(y, x) \\
\Longleftrightarrow &\exp\left(\theta f_t(x) + \theta f_{t+1}(y)\right) \leq \exp\left(\theta f_t(y) + \theta f_{t+1}(x)\right) \\
\Longleftrightarrow &f_t(x) - f_t(y) \leq f_{t+1}(x) - f_{t+1}(y)
\end{aligned}$$

The last line holds $\forall x \leq y$ due to the submodularity of f_t. Consequently, we can use Theorem 1.C.20 to prove Eq. (3.8). $\square$

Proof of Theorem 3.1

Proof. 1. Define $g_t(\theta, s_t) = U_t(\theta, s_t) - \mu_\theta\left(\tilde{G}(s_t)\right)$. The optimal policy dictates ceasing treatment if $g_t(\theta, s_t) = 0$ and continuing otherwise. Then, according to the backward recursion of $U_t(\theta, s_t)$, we can establish the corresponding recursion of $g_t(\theta, s)$: for any $t \in \mathcal{T}/\{T\}$,

$$g_t(\theta, s) = \min\left\{0, H(s) + \frac{1}{\theta}\log\left(\sum_{s'} P(s, s')\exp\left(\theta \cdot \left(g_{t+1}(\theta, s') + \mu_\theta\left(\tilde{G}(s')\right)\right)\right)\right) - \mu_\theta\left(\tilde{G}(s)\right)\right\}.$$

We use mathematical induction to prove that $g_t(\theta, s)$ is increasing in s, which implies the threshold structure of the policy. $g_T(\theta, s) = 0$. Based on Assumption 3.3,

$$g_{T-1}(\theta, s) = \min\left\{0, H(s) + \frac{1}{\theta}\log\left(\sum_{s'} P(s, s')\exp\left(\theta \cdot \mu_\theta\left(\tilde{G}(s')\right)\right)\right) - \mu_\theta\left(\tilde{G}(s)\right)\right\}$$

is increasing in s. Then, assume the statement holds for $t = k$. Consider $t = k - 1$. The key idea is to prove that: for any $s \geq s^\dagger$,

$$\frac{\sum_{s'} P(s, s')\exp\left(\theta g_k(\theta, s') + \theta\mu_\theta\left(\tilde{G}(s')\right)\right)}{\sum_{s'} P(s^\dagger, s')\exp\left(\theta g_k(\theta, s') + \theta\mu_\theta\left(\tilde{G}(s')\right)\right)} \geq \frac{\sum_{s'} P(s, s')\exp\left(\theta\mu_\theta\left(\tilde{G}(s')\right)\right)}{\sum_{s'} P(s^\dagger, s')\exp\left(\theta\mu_\theta\left(\tilde{G}(s')\right)\right)} \tag{A.18}$$

The proof is based on Theorem 1.B.48. from Shaked and Shanthikumar (2007). Notably, MRL order implies the reserve hazard rate order, *i.e.*, $\tilde{x} \leq_{lr} \tilde{y} \implies \tilde{x} \leq_{rh} \tilde{y}$.

Theorem 1.B.48: Let $\tilde{x}$ and $\tilde{y}$ be two independent random variables. Then $\tilde{x} \leq_{rh} \tilde{y}$ if, and only if,

$$\mathbb{E}_{\mathbb{P}}\left[\phi_1(\tilde{x}, \tilde{y})\right] \leq \mathbb{E}_{\mathbb{P}}\left[\phi_2(\tilde{x}, \tilde{y})\right]$$

for all ϕ_1 and ϕ_2 which satisfy that, for each y, $\Delta\phi_{21}(x, y)$ decreases in x on $\{x \leq y\}$; and $\Delta\phi_{21}(x, y) \geq -\Delta\phi_{21}(y, x)$ whenever $x \leq y$, where $\Delta\phi_{21}(x, y) := \phi_2(x, y) - \phi_1(x, y)$.

Define $\begin{cases} \phi_1(x, y) = \exp\left(\theta\mu_\theta\left(\tilde{G}(y)\right) + \theta g_k(\theta, x) + \theta\mu_\theta\left(\tilde{G}(x)\right)\right), \\ \phi_2(x, y) = \exp\left(\theta\mu_\theta\left(\tilde{G}(y)\right) + \theta g_k(\theta, y) + \theta\mu_\theta\left(\tilde{G}(x)\right)\right). \end{cases}$ Then

$$\Delta\phi_{21}(x, y) = \exp\left(\theta\mu_\theta\left(\tilde{G}(x)\right) + \theta\mu_\theta\left(\tilde{G}(y)\right)\right) \cdot \left(\exp\left(\theta g_k(\theta, y)\right) - \exp\left(\theta g_k(\theta, x)\right)\right),$$

is decreasing in x on $\{x \leq y\}$ based on the assumption when $t = k$. Also, $\Delta\phi_{21}(x, y) = -\Delta\phi_{21}(y, x)$. These two functions satisfy the conditions of Theorem 1.B.48. Let $\tilde{y} \sim P_s$, and

$\tilde{x} \sim P_{s^\dagger}$. We have $\tilde{x} \leq_{lr} \tilde{y}$. Then based on Theorem 1.B.48, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\mu_\theta\left(\tilde{G}(\tilde{y})\right)+\theta g_k(\theta,\tilde{y})\right)\right]\cdot\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\mu_\theta\left(\tilde{G}(\tilde{x})\right)\right)\right]\\
\geq&\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\mu_\theta\left(\tilde{G}(\tilde{x})\right)+\theta g_k(\theta,\tilde{x})\right)\right]\cdot\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\mu_\theta\left(\tilde{G}(\tilde{y})\right)\right)\right],
\end{aligned}$$

which is equivalent to Eq. (A.18). Using Eq. (A.18), we can show that

$$\begin{aligned}
&H(s)+\frac{1}{\theta}\log\left(\sum_{s'}P(s,s')\exp\left(\theta\cdot\mu_\theta\left(\tilde{G}(s')\right)+\theta\cdot g_k(\theta,s')\right)\right)-\mu_\theta\left(\tilde{G}(s)\right)\\
&-\left[H(s^\dagger)+\frac{1}{\theta}\log\left(\sum_{s'}P(s^\dagger,s')\exp\left(\theta\cdot\mu_\theta\left(\tilde{G}(s')\right)+\theta\cdot g_k(\theta,s')\right)\right)-\mu_\theta\left(\tilde{G}(s^\dagger)\right)\right]\\
&\geq H(s)+\frac{1}{\theta}\log\left(\sum_{s'}P(s,s')\exp\left(\theta\cdot\mu_\theta\left(\tilde{G}(s')\right)\right)\right)-\mu_\theta\left(\tilde{G}(s)\right)\\
&-\left(H(s^\dagger)+\frac{1}{\theta}\log\left(\sum_{s'}P(s^\dagger,s')\exp\left(\theta\cdot\mu_\theta\left(\tilde{G}(s')\right)\right)\right)-\mu_\theta\left(\tilde{G}(s^\dagger)\right)\right)\\
&\geq 0.
\end{aligned}$$

The second inequality is due to Assumption 3.3. Thus, $g_{k-1}(\theta,s)$ is increasing in s. By mathematical induction, we prove that $U_t(\theta,s)-\mu_\theta\left(\tilde{G}(s)\right)$ is increasing in s for any t. Therefore, a threshold structure exists.

The properties of subsets $\mathcal{R}^0_{t,\theta}$ and $\mathcal{R}^1_{t,\theta}$ are quick results according to Proposition 3.4.

2. The submodularity of U function is equivalent to that for any $t\in\mathcal{T}\backslash\{T\}$, $s_1<s_2\in\mathcal{S}$, $U_{t+1}(\theta,s_1)-U_t(\theta,s_1)\geq U_{t+1}(\theta,s_2)-U_t(\theta,s_2)$. In the following analysis, we use mathematical induction in index t to prove it. When $t=T-1$, it holds based on Assumption 3.3.

Assume the statement holds when $t\geq\tau$. Then, let us consider one period backward, $t=\tau-1$. We need to show that

$$\begin{aligned}
&U_\tau(\theta,s_1)-U_{\tau-1}(\theta,s_1)\geq U_\tau(\theta,s_2)-U_{\tau-1}(\theta,s_2)\\
\Longleftrightarrow&U_\tau(\theta,s_1)-U_\tau(\theta,s_2)\geq U_{\tau-1}(\theta,s_1)-U_{\tau-1}(\theta,s_2),
\end{aligned}\tag{A.19}$$

where $U_\tau(\theta,s_1)=\min\left\{\mu_\theta\left(\tilde{G}(s_1)\right),H(s_1)+\frac{1}{\theta}\log\sum_{s\in\mathcal{S}}P(s_1,s)\exp\left(\theta\cdot U_{\tau+1}(\theta,s)\right)\right\}$. Then, we check Eq. A.19 using minimal values in different scenarios. Due to the threshold structure of policies, the action matrix $\begin{pmatrix}\pi^*_{\tau-1}(\theta,s_2) & \pi^*_\tau(\theta,s_2)\\ \pi^*_{\tau-1}(\theta,s_1) & \pi^*_\tau(\theta,s_1)\end{pmatrix}$ can take a value from $\begin{pmatrix}0&0\\0&0\end{pmatrix}$, $\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $\begin{pmatrix}0&1\\0&1\end{pmatrix}$, $\begin{pmatrix}1&1\\0&0\end{pmatrix}$, $\begin{pmatrix}1&1\\0&1\end{pmatrix}$ or $\begin{pmatrix}1&1\\1&1\end{pmatrix}$.

- $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Assumption at $t = \tau$ guarantees the condition of Proposition 3.5, which shows that

$$
\begin{aligned}
&H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) \\
\geq &H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau}(\theta, s)\right)
\end{aligned}
$$

- $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. We can show that

$$
\begin{aligned}
&H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) - \mu_\theta\left(\tilde{G}(s_2)\right) \\
\geq &H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) \\
\geq &H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau}(\theta, s)\right),
\end{aligned}
$$

where the first inequality is due to the fact that $\pi_\tau^{\theta\star}(s_2) = 1$, i.e.

$$
\mu_\theta(\tilde{G}(s_2)) < H(s_2) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right),
$$

and in the second inequality we use the result of Proposition 3.5 again.

- $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$. Since we assume statements hold for any later periods ($t \geq \tau$), we have that

$$
\begin{aligned}
&H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau}(\theta, s)\right) \\
\leq &H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{\tau+1}(\theta, s)\right) \\
\leq &\ldots \\
\leq &H(s_1) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1, s)\exp\left(\theta \cdot U_{T}(\theta, s)\right) - H(s_2) - \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_2, s)\exp\left(\theta \cdot U_{T}(\theta, s)\right) \\
\leq &U_T(\theta, s_1) - U_T(\theta, s_2) = \mu_\theta\left(\tilde{G}(s_1)\right) - \mu_\theta\left(\tilde{G}(s_2)\right).
\end{aligned}
$$

In the inequalities before the last line, we use both the induction assumptions and Proposition 3.5. The induction assumption at time t implies that $U_{t+1}(\theta, s_1) - U_{t+1}(\theta, s_2) \geq$

$U_t(\theta, s_1) - U_t(\theta, s_2)$, while Proposition 3.5 demonstrates

$$\begin{aligned}
&\frac{1}{\theta}\log\left(\sum_{s\in\mathcal{S}} P(s_1,s)\exp\left(U_{t+1}(\theta,s)\right)\right) - \frac{1}{\theta}\log\left(\sum_{s\in\mathcal{S}} P(s_2,s)\exp\left(U_{t+1}(\theta,s)\right)\right) \\
&\geq \frac{1}{\theta}\log\left(\sum_{s\in\mathcal{S}} P(s_1,s)\exp\left(U_{t}(\theta,s)\right)\right) - \frac{1}{\theta}\log\left(\sum_{s\in\mathcal{S}} P(s_2,s)\exp\left(U_{t}(\theta,s)\right)\right).
\end{aligned}$$

Then, we can conduct such an evaluation forward until reaching the last period T. The last inequality is based on Assumption 3.3.

- $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. It holds due to the monotonicity of U function in t, *i.e.*,

$$\frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_1,s)\exp\left(\theta\cdot U_{\tau+1}(\theta,s)\right) \geq \log\sum_{s\in\mathcal{S}} P(s_1,s)\exp\left(\theta\cdot U_{\tau}(\theta,s)\right).$$

- $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ are trival cases.

By mathematical induction, we prove that $U_t(\theta, s)$ is submodular in $\mathcal{T} \times \mathcal{S}$. □

Proof of Theorem 3.2

Proof. First, we have

$$\begin{aligned}
&H(s_t) + \frac{1}{\theta}\cdot\log\sum_{s_{t+1}} P(s_t,s_{t+1})\cdot\exp\left(\theta\cdot\mu_\theta\left(\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right)\right) \\
&= H(s_t) + \frac{1}{\theta}\cdot\log\sum_{s_{t+1}} P(s_t,s_{t+1})\exp\left(\theta\cdot\frac{1}{\theta}\log\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\cdot\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right)\right]\right) \\
&= H(s_t) + \frac{1}{\theta}\cdot\log\sum_{s_{t+1}} P(s_t,s_{t+1})\left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\cdot\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right)\right]\right] \\
&\geq H(s_t) + \frac{1}{\theta}\cdot\sum_{s_{t+1}} P(s_t,s_{t+1})\log\left[\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta\cdot\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right)\right]\right] \\
&\geq H(s_t) + \frac{1}{\theta}\cdot\sum_{s_{t+1}} P(s,s_{t+1})\left[\mathbb{E}_{\mathbb{P}}\left[\log\left[\exp\left(\theta\cdot\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right)\right]\right]\right] \\
&= H(s_t) + \sum_{s_{t+1}} P(s_t,s_{t+1})\mathbb{E}_{\mathbb{P}}\left[\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right] \\
&\geq H(s_t) + \sum_{s_{t+1}} P(s_t,s_{t+1})\mathbb{E}_{\mathbb{P}}\left[\tilde{V}_{t+1}^{\boldsymbol{\pi}_0^\star}(s_{t+1})\right],
\end{aligned} \tag{A.20}$$

where the first two inequalities are based on Jensen's inequality, and the last inequality is due to the fact that $\boldsymbol{\pi}_0^\star$ optimizes the expected cost-to-go. For the ROSP, the benefit function is

$$h_t^\theta(s_t) = G(s_t) - H(s_t) - \frac{1}{\theta} \cdot \log \sum_{s_{t+1}} P(s_t, s_{t+1}) \cdot \exp\left(\theta \cdot \mu_\theta\left(\tilde{V}_{t+1}^{\boldsymbol{\pi}_\theta^\star}(s_{t+1})\right)\right)$$

while for NOSP, the benefit function is

$$h_t^0(s_t) = G(s_t) - H(s_t) - \sum_{s_{t+1}} P(s_t, s_{t+1}) \mathbb{E}_{\mathbb{P}}\left[\tilde{V}_{t+1}^{\boldsymbol{\pi}_0^\star}(s_{t+1})\right].$$

Then we can get that the benefit function of the ROSP is always smaller than or equal to the benefit of the NOSP, *i.e.*, $h_t^\theta(s) \leq h_t^0(s), \forall s \in \mathcal{S}, t \in \mathcal{T}$. Consequently, $\{s \,|\, h_t^0(s) \leq 0\} \subseteq \{s \,|\, h_t^\theta(s) \leq 0\}$, which implies the optimal policy of ROSP is more proactive than the optimal policy of NOSP. It is easy to show that in this case $h_t^\theta(s)$ is decreasing w.r.t. θ. Therefore, the optimal treatment policy of ROSP is more aggressive given a larger θ. □

Proof of Lemma 3.1

Proof. Define a function, $f(\boldsymbol{x}) = \frac{1}{\theta} \log \sum_{s' \in \mathcal{S}} P(s, s') \exp(\theta \cdot x_{s'})$, which is known as the Log-SumExp function. $f(\boldsymbol{x})$ is convex in $\boldsymbol{x}$ (Boyd and Vandenberghe, 2004). We have

$$\begin{aligned}
h_t^\theta(s) - h_{t+1}^\theta(s) &= -\frac{1}{\theta} \log \sum_{s' \in \mathcal{S}} P(s, s') \exp\left(\theta U_{t+1}(\theta, s')\right) + \frac{1}{\theta} \log \sum_{s' \in \mathcal{S}} P(s, s') \exp\left(\theta U_{t+2}(\theta, s')\right) \\
&\geq \left(\nabla f(U_{t+1}(\theta, 1), ..., U_{t+1}(\theta, S))\right)^\top \begin{pmatrix} U_{t+2}(\theta, 1) - U_{t+1}(\theta, 1) \\ ... \\ U_{t+2}(\theta, S) - U_{t+1}(\theta, S) \end{pmatrix} \\
&= \sum_{c \in \mathcal{S}} \frac{P(s, c) \exp\left(\theta U_{t+1}(\theta, c)\right) \left(U_{t+2}(\theta, c) - U_{t+1}(\theta, c)\right)}{\sum_{s' \in \mathcal{S}} P(s, s') \exp\left(\theta U_{t+1}(\theta, s')\right)} \\
&\geq \sum_{s' \in \mathcal{S}} P(s, s')(U_{t+2}(\theta, s') - U_{t+1}(\theta, s')). \qquad \text{(A.21)}
\end{aligned}$$

The first inequality is due to the fact that $f(\boldsymbol{x}) - f(\boldsymbol{y}) \geq \nabla f(\boldsymbol{y})^\top (\boldsymbol{x} - \boldsymbol{y})$ since $f(\boldsymbol{x})$ is convex. In the following, we aim to prove the second inequality.

Considering a probability mass function $\hat{\boldsymbol{P}}_s$, where

$$\hat{P}_s(c) = \frac{P(s, c) \exp\left(\theta U_{t+1}(\theta, c)\right)}{\sum_{s' \in \mathcal{S}} P(s, s') \exp\left(\theta U_{t+1}(\theta, s')\right)},$$

we have the following relationship $\hat{\boldsymbol{P}}_s \leq_{st} \boldsymbol{P}_s$, since $U_{t+1}(\theta, c)$ is decreasing in c. Based on the property of FSD and the submodularity shown in Theorem 3.1, that is $U_{t+2}(\theta, c) - U_{t+1}(\theta, c)$ is decreasing in c, we prove the second inequality in (A.21). Therefore, we complete the proof of this lemma. $\square$

Proof of Theorem 3.3

Proof. To investigate when ROSP provides more conservative policies than NOSP does, it suffices to show, for any $t \leq T-1$,

$$h_t^\theta(s) \geq h_t^0(s),\ \forall s \in \mathcal{R}_{0,0}^0. \tag{A.22}$$

We use mathematical induction to prove the above statement. To do that, we require the following additional equation to ensure the induction step.

$$U_t(\theta, s) - U_{t+1}(\theta, s) \leq U_t(0, s) - U_{t+1}(0, s),\ \forall s \in \mathcal{S},\ t \leq T-1. \tag{A.23}$$

Based on our assumptions, we have that $h_{T-1}^\theta(s) \geq h_{T-1}^0(s),\ \forall s \in \mathcal{R}_{0,0}^0$. In the last period, Eq. (A.23) is equivalent to

$$\min\left\{\frac{1}{\theta}\sum_{s'\in\mathcal{S}} P(s,s')\exp\left(\theta\cdot\mu_\theta\left(\tilde{G}(s')\right)\right), \mu_\theta\left(\tilde{G}(s)\right)\right\} - \mu_\theta\left(\tilde{G}(s)\right) \leq \min\left\{\sum_{s'\in\mathcal{S}} P(s,s')\mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s')\right], \mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s)\right]\right\} - \mathbb{E}_{\mathbb{P}}\left[\tilde{G}(s)\right]$$

$$\Longleftrightarrow \max\left\{h_{T-1}^\theta(s), 0\right\} \geq \max\left\{h_{T-1}^0(s), 0\right\}$$

Proposition 3.4 guarantees that $\mathcal{R}_{T-1,0}^0 \subset \mathcal{R}_{0,0}^0$. Therefore, $h_{T-1}^\theta(s) \geq h_{T-1}^0(s) = \max\left\{h_{T-1}^0(s), 0\right\}$ for any $s \in \mathcal{R}_{T-1,0}^0$, while $\max\left\{h_{T-1}^\theta(s), 0\right\} \geq \max\left\{h_{T-1}^0(s), 0\right\} = 0$ for any $s \notin \mathcal{R}_{T-1,0}^0$. Eq. (A.23) holds at time $T-1$.

Assume Eqs. (A.22) and (A.23) hold for any $t \geq k$. Then we consider the case when $t = k-1$.

$$\begin{aligned} h_{k-1}^\theta(s) - h_k^\theta(s) &\geq \sum_{s'\in\mathcal{S}} P(s,s')(U_{k+1}(\theta,s') - U_k(\theta,s')) \\ &\geq \sum_{s'\in\mathcal{S}} P(s,s')(U_{k+1}(0,s') - U_k(0,s')) \\ &= h_{k-1}^0(s) - h_k^0(s), \end{aligned} \tag{A.24}$$

where the first inequality is based on Lemma 3.1 and the second inequality is due to our induction assumption. Since $h_k^\theta(s) \geq h_k^0(s),\ \forall s \in \mathcal{R}_{0,0}^0$, we have $h_{k-1}^\theta(s) \geq h_{k-1}^0(s),\ \forall s \in \mathcal{R}_{0,0}^0$. Therefore, Eq. (A.22) holds at $k-1$. Then, we consider Eq. (A.23) at time $k-1$ that is equivalent to

$$\max\{h_{k-1}^\theta(s), 0\} - \max\{h_k^\theta(s), 0\} \geq \max\{h_{k-1}^0(s), 0\} - \max\{h_k^0(s), 0\}. \tag{A.25}$$

For any s, we need to consider the following scenarios:

- If $h_k^0(s) > 0$, it follows that $h_{k-1}^0(s) > 0$ since $h_k^0(s)$ decreases in k. Because $s \in \mathcal{R}_{k,0}^0 \subset \mathcal{R}_{0,0}^0$, Eq. (A.22) at $t = k$ implies that $h_{k-1}^\theta(s) \geq h_k^\theta(s) \geq h_k^0(s) \geq 0$. Then, based on Eq. (A.24), we prove that Eq. (A.25).

- If $h_k^0(s) \leq 0$, $h_{k-1}^0(s)$ can be either negative or positive.
  - If $h_{k-1}^0(s) > 0$, then $h_{k-1}^\theta(s) > 0$. We need to consider the following two cases:
    * If $h_k^\theta(s) > 0$, Eq. (A.24) implies $h_{k-1}^\theta(s) - h_k^\theta(s) \geq h_{k-1}^0(s) - h_k^0(s) \geq h_{k-1}^0(s) - 0$. Thus, Eq. (A.25) holds.
    * If $h_k^\theta(s) \leq 0$, Eq. (A.25) is equivalent to $h_{k-1}^\theta(s) - 0 \geq h_{k-1}^0(s) - 0$. As $h_{k-1}^0(s) \geq 0$, $s \in \mathcal{R}_{k-1,0}^0 \subset \mathcal{R}_{0,0}^0$. Eq. (A.22) at time $k-1$ shows that $h_{k-1}^\theta(s) \geq h_{k-1}^0(s)$. Thus, Eq. (A.25) holds.
  - If $h_{k-1}^0(s) \leq 0$, then $\max\{h_{k-1}^0(s), 0\} - \max\{h_k^0(s), 0\} = 0$. Additionally, since $h_k^\theta(s)$ decreases in k, it follows that $\max\{h_{k-1}^\theta(s), 0\} - \max\{h_k^\theta(s), 0\} \geq 0$. Thus, Eq. (A.25) holds.

This shows that $U_{k-1}(\theta, s) - U_k(\theta, s) \leq U_{k-1}(0, s) - U_k(0, s)$. By mathematical induction, we prove both Eqs. (A.22) and (A.23), while Eq. (A.22) implies that ROSP is more conservative than NOSP under the given condition. $\square$

Proof of Proposition 3.6

Proof. This Proposition is proved by mathematical induction backward recursively. For simplicity, we omit the arguments for the expectation, i.e. $\theta = 0$. The detail can be found in Cheng, et al. (2024). First, when $t = T$,

$$
\begin{aligned}
\bar{U}_T^P(\theta, s_T) &= \frac{1}{\theta} \log \sum_{\hat{s} \in \mathcal{S}} \tilde{Q}(s_T, \hat{s}) \exp\left(\theta \cdot U_T^P(\theta, s_T, \hat{s})\right) \\
&= \frac{1}{\theta} \log \sum_{\hat{s} \in \mathcal{S}} \tilde{Q}(s_T, \hat{s}) \exp\left(\theta \cdot \mu_\theta\left(\tilde{G}(s_T)\right)\right) \\
&= \mu_\theta\left(\tilde{G}(s_T)\right).
\end{aligned}
$$

Thus, the statement holds at time T, *i.e.*, $\bar{U}_T^P(\theta, s_T) \leq U_T(\theta, s_T)$, $\forall s_T \in \mathcal{S}$. Assume the statement is true for epoch $t+1$, *i.e.*, $\bar{U}_{t+1}^P(\theta, s_{t+1}) \leq U_{t+1}(\theta, s_{t+1})$, $\forall s_{t+1} \in \mathcal{S}$. For epoch t, we have

$$
\begin{aligned}
&U_t(\theta, s_t) = \min\left\{\mu_\theta\left(\tilde{G}(s_t)\right), H(s_t) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_t, s)\exp\left(\theta\cdot U_{t+1}(\theta, s)\right)\right\}\\
&\geq \min\left\{\mu_\theta\left(\tilde{G}(s_t)\right), H(s_t) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_t, s)\exp\left(\theta\cdot \bar{U}^P_{t+1}(\theta, s)\right)\right\}\\
&= \min\left\{\mu_\theta\left(\tilde{G}(s_t)\right), H(s_t) + \frac{1}{\theta}\log\sum_{s\in\mathcal{S}} P(s_t, s)\sum_{\hat{s}\in\mathcal{S}}\tilde{Q}(s,\hat{s})\exp\left(\theta\cdot U^P_{t+1}(\theta, s, \hat{s})\right)\right\}\\
&= \min\left\{\mu_\theta\left(\tilde{G}(s_t)\right), H(s_t) + \frac{1}{\theta}\log\sum_{\hat{s}_{t+1}\in\mathcal{S}} \tilde{Q}(s_t, \hat{s}_{t+1})\sum_{s\in\mathcal{S}} q(s|s_t,\hat{s}_{t+1})\sum_{\hat{s}\in\mathcal{S}}\tilde{Q}(s,\hat{s})\exp\left(\theta\cdot U^P_{t+1}(\theta, s, \hat{s})\right)\right\}\\
&\geq \frac{1}{\theta}\log\sum_{\hat{s}_{t+1}\in\mathcal{S}} \tilde{Q}(s_t, \hat{s}_{t+1})\\
&\qquad \exp\left(\theta\cdot\min\left\{\mu_\theta\left(\tilde{G}(s_t)\right), H(s_t) + \frac{1}{\theta}\log\sum_{s,\hat{s}\in\mathcal{S}} q(s|s_t,\hat{s}_{t+1})\tilde{Q}(s,\hat{s})\exp\left(\theta\cdot U^P_{t+1}(\theta, s, \hat{s})\right)\right\}\right)\\
&= \frac{1}{\theta}\log\sum_{\hat{s}_{t+1}\in\mathcal{S}} \tilde{Q}(s_t, \hat{s}_{t+1})\exp\left(\theta\cdot U^P_t(\theta, s_t, \hat{s}_{t+1})\right)\\
&= \bar{U}^P_t(\theta, s_t).
\end{aligned}
$$

The first inequality is a result of the induction assumption, and Jensen's inequality is applied to the minimum in the second inequality. The first line and the seventh line use the optimality equation. The last line follows from the definition of $\bar{U}^P_t(\theta, s_t)$. Thus, by mathematical induction, Proposition 3.6 holds. □

Proof of Proposition 3.7

Proof. Proof. We prove the monotonicity of $U^{\mathrm{P}}_t(\theta, s_t, \hat{s}_{t+1})$ using mathematical induction. The proof relies on the following lemma about the conditional transition matrices.

Lemma A.2. *The conditional transition matrices* $\left[q(s_{t+1}|s_t, \hat{s}_{t+1})\right]_{s_{t+1}, s_t\in\mathcal{S}}$ *and* $\left[q(s_{t+1}|s_t, \hat{s}_{t+1})\right]_{s_{t+1}, \hat{s}_{t+1}\in\mathcal{S}}$ *are TP2.*

Proof. Proof. of Lemma A.2 By definition, the conditional transition probability is:

$$
q(s_{t+1}|s_t, \hat{s}_{t+1}) = \frac{P(s_t, s_{t+1})Q(s_{t+1}, \hat{s}_{t+1})}{\sum_{s\in\mathcal{S}} P(s_t, s)Q(s, \hat{s}_{t+1})}.
$$

We first prove the TP2 property for the first matrix. According to the definition of TP2, it is equivalent to showcasing the second minor is non-negative, *i.e.*, $\forall s_{t+1} \leq s^{\dagger}_{t+1}, s_t \leq s^{\dagger}_t$,

$$\begin{aligned} & q(s_{t+1}|s_t, \hat{s}_{t+1})q(s^{\dagger}_{t+1}|s^{\dagger}_t, \hat{s}_{t+1}) - q(s^{\dagger}_{t+1}|s_t, \hat{s}_{t+1})q(s_{t+1}|s^{\dagger}_t, \hat{s}_{t+1}) \\ = & \tfrac{P(s_t,s_{t+1})Q(s_{t+1},\hat{s}_{t+1})P(s^{\dagger}_t,s^{\dagger}_{t+1})Q(s^{\dagger}_{t+1},\hat{s}_{t+1})}{\tilde{Q}(s_t,\hat{s}_{t+1})\tilde{Q}(s^{\dagger}_t,\hat{s}_{t+1})} + \tfrac{P(s_t,s^{\dagger}_{t+1})Q(s^{\dagger}_{t+1},\hat{s}_{t+1})P(s^{\dagger}_t,s_{t+1})Q(s_{t+1},\hat{s}_{t+1})}{\tilde{Q}(s_t,\hat{s}_{t+1})\tilde{Q}(s^{\dagger}_t,\hat{s}_{t+1})} \\ = & \tfrac{Q(s_{t+1},\hat{s}_{t+1})Q(s^{\dagger}_{t+1},\hat{s}_{t+1})}{\tilde{Q}(s_t,\hat{s}_{t+1})\tilde{Q}(s^{\dagger}_t,\hat{s}_{t+1})}(P(s_t, s_{t+1})P(s^{\dagger}_t, s^{\dagger}_{t+1}) - P(s^{\dagger}_t, s_{t+1})P(s_t, s^{\dagger}_{t+1})) \geq 0. \end{aligned} \tag{A.26}$$

The last inequality is based on the fact that $\mathbf{P}$ is TP2. Then, we consider the second matrix. According to the definition TP2, it is equivalent to consider the second minor such that $\forall s_{t+1} \leq s^{\dagger}_{t+1}, \hat{s}_{t+1} \leq \hat{s}^{\dagger}_{t+1}$,

$$\begin{aligned} & q(s_{t+1}|s_t, \hat{s}_{t+1})q(s^{\dagger}_{t+1}|s_t, \hat{s}^{\dagger}_{t+1}) - q(s^{\dagger}_{t+1}|s_t, \hat{s}_{t+1})q(s_{t+1}|s_t, \hat{s}^{\dagger}_{t+1}) \\ = & \tfrac{P(s_t,s_{t+1})Q(s_{t+1},\hat{s}_{t+1})P(s_t,s^{\dagger}_{t+1})Q(s^{\dagger}_{t+1},\hat{s}^{\dagger}_{t+1})}{\tilde{Q}(s_t,\hat{s}_{t+1})\tilde{Q}(s_t,\hat{s}^{\dagger}_{t+1})} + \tfrac{P(s_t,s^{\dagger}_{t+1})Q(s^{\dagger}_{t+1},\hat{s}_{t+1})P(s_t,s_{t+1})Q(s_{t+1},\hat{s}^{\dagger}_{t+1})}{\tilde{Q}(s_t,\hat{s}_{t+1})\tilde{Q}(s_t,\hat{s}^{\dagger}_{t+1})} \\ = & \tfrac{P(s_t,s_{t+1}),P(s_t,s^{\dagger}_{t+1})}{\tilde{Q}(s_t,\hat{s}_{t+1})\tilde{Q}(s_t,\hat{s}^{\dagger}_{t+1})}(Q(s_{t+1}, \hat{s}_{t+1})Q(s^{\dagger}_{t+1}, \hat{s}^{\dagger}_{t+1}) - Q(s^{\dagger}_{t+1}, \hat{s}_{t+1})Q(s_{t+1}, \hat{s}^{\dagger}_{t+1})) \geq 0. \end{aligned} \tag{A.27}$$

The last inequality is based on the fact that $\mathbf{Q}$ is TP2. Therefore, we prove both two conditional transition matrices are TP2. $\square$

Then, we use backward induction on t to prove the monotonicity of $U^P_t(\theta, s_t, \hat{s}_{t+1})$. Recall that for any $t \in \mathcal{T}\backslash\{T\}$ and $\theta > 0$, the Bellman recursion for ROSP-P is:

$$U^{\mathrm{P}}_t(\theta, s_t, \hat{s}_{t+1}) \;=\; \min\left\{ \mu_\theta\big(\tilde{G}(s_t)\big),\, H(s_t) + \tfrac{1}{\theta}\log \textstyle\sum_{s_{t+1},\,\hat{s}_{t+2}\in\mathcal{S}} \mathbb{P}\big(s_{t+1}, \hat{s}_{t+2} \,\big|\, s_t, \hat{s}_{t+1}\big)\, e^{\theta\, U^{\mathrm{P}}_{t+1}(\theta, s_{t+1}, \hat{s}_{t+2})} \right\},$$

where $\mathbb{P}(s_{t+1}, \hat{s}_{t+2} \mid s_t, \hat{s}_{t+1}) = q(s_{t+1} \mid s_t, \hat{s}_{t+1})\, \tilde{Q}(s_{t+1}, \hat{s}_{t+2})$. For $\theta = 0$, the exponential term reduces to an expectation, but the proof is analogous. In the following, we focus on the case where $\theta > 0$.

Step 1 (Base Case). At $t = T$,

$$U^{\mathrm{P}}_T(\theta, s_T, \hat{s}_{T+1}) \;=\; \mu_\theta\big(\tilde{G}(s_T)\big).$$

By Assumption 3.1, if $s_1 \geq s_2$, then $\tilde{G}(s_1) \leq_{st} \tilde{G}(s_2)$ and hence $\mu_\theta\big(\tilde{G}(s_1)\big) \;\leq\; \mu_\theta\big(\tilde{G}(s_2)\big)$ for all $\theta \geq 0$. Thus, for any $s_1 \leq s_2$ and $\hat{s}_1 \leq \hat{s}_2$,

$$U^{\mathrm{P}}_T(\theta, s_1, \hat{s}_{T+1}) \;\geq\; U^{\mathrm{P}}_T(\theta, s_2, \hat{s}_{T+1}) \quad \text{and} \quad U^{\mathrm{P}}_T(\theta, s_T, \hat{s}_1) \;=\; U^{\mathrm{P}}_T(\theta, s_T, \hat{s}_2)$$

Therefore, $U^{\mathrm{P}}_T(\theta, \cdot, \cdot)$ is decreasing in s_T and $\hat{s}_{T+1}$.

Step 2 (Inductive Step). Consider the time t and assume $U^{\mathrm{P}}_{t+1}(\theta, s_{t+1}, \hat{s}_{t+2})$ is decreasing in s_{t+1}, $\hat{s}_{t+2}$. We next show $U^{\mathrm{P}}_{t}(\theta, s_t, \hat{s}_{t+1})$ is also decreasing in s_t, $\hat{s}_{t+1}$. We consider the two terms within minimization operator:

- **Monotonicity of termination cost.** By Assumption 3.1, $\mu_\theta\big(\tilde{G}(s)\big)$ is decreasing in s.
- **Monotonicity of continuation cost.** By Assumption 3.1, $H(s)$ is decreasing in s. Define

$$\Phi\big[s_t, \hat{s}_{t+1}\big] \;:=\; \frac{1}{\theta}\log \sum_{s_{t+1},\, \hat{s}_{t+2}} \mathbb{P}\big(s_{t+1}, \hat{s}_{t+2} \,\big|\, s_t, \hat{s}_{t+1}\big)\; \exp\Big(\theta\, U^{\mathrm{P}}_{t+1}(\theta, s_{t+1}, \hat{s}_{t+2})\Big).$$

We wish to show $\Phi[s_t, \hat{s}_{t+1}]$ is nonincreasing in s_t and $\hat{s}_{t+1}$. Since $U^{\mathrm{P}}_{t+1}(\theta, \cdot, \cdot)$ is decreasing in its state and predicted state by the induction hypothesis, we need the transition kernel $\mathbb{P}\big(s_{t+1}, \hat{s}_{t+2} \,\big|\, s_t, \hat{s}_{t+1}\big)$ to preserve this monotonic relationship in the log-sum-exp operator. Observe that

$$\mathbb{P}\big(s_{t+1}, \hat{s}_{t+2} \,\big|\, s_t, \hat{s}_{t+1}\big) \;=\; q(s_{t+1} \mid s_t, \hat{s}_{t+1})\, \tilde{Q}\big(s_{t+1}, \hat{s}_{t+2}\big),$$

By Assumption 3.2 (and Q also being TP2), the product matrix $\tilde{\mathbf{Q}}$ is TP2. Thus,

$$\sum_{\hat{s}_{t+2}} \tilde{Q}(s_{t+1}, \hat{s}_{t+2}) \exp\Big(\theta\, U^{\mathrm{P}}_{t+1}(\theta, s_{t+1}, \hat{s}_{t+2})\Big)$$

is decreasing in s_{t+1}. From Lemma A.2, we know that $\big[q(s_{t+1}|s_t, \hat{s}_{t+1})\big]_{s_{t+1}, s_t \in \mathcal{S}}$ and $\big[q(s_{t+1}|s_t, \hat{s}_{t+1})\big]_{s_{t+1}, \hat{s}_{t+1} \in \mathcal{S}}$ are TP2. Therefore, as s_t or $\hat{s}_{t+1}$ increases, the expectation (or summation) w.r.t. the conditional transition distribution preserves the ordering. Furthermore, the logarithm function $u \;\mapsto\; \frac{1}{\theta}\log\Big[u\Big]$ is strictly increasing in each u. Hence, combining these two facts, we deduce that $\Phi[s_t, \hat{s}_{t+1}]$ is decreasing in s_t.

Since both two terms in the minimization are decreasing, we conclude that $U^{\mathrm{P}}_{t}(\theta, s_t, \hat{s}_{t+1})$ is decreasing in s_t and $\hat{s}_{t+1}$.

Step 3 (Final Conclusion). By induction on t from T down to 0, we conclude that $U^{\mathrm{P}}_{t}(\theta, s_t, \hat{s}_{t+1})$ is decreasing in s_t and $\hat{s}_{t+1}$ for all $t \in \mathcal{T}$ and all $\theta \geq 0$. □

Proof of Proposition 3.8

Proof. Proof. In the proof of Proposition 3.7, we have already shown that the log-sum-exp term

$$\frac{1}{\theta}\log \sum_{s', \hat{s}' \in \mathcal{S}} \mathbb{P}(s', \hat{s}' | s, \hat{s}) \exp\left(\theta \cdot U^{\mathrm{P}}_{t+1}(\theta, s', \hat{s}')\right)$$

is decreasing in $\hat{s}$. Thus, $h_t^{\mathrm{P},\theta}(s,\hat{s})$ is increasing in $\hat{s}$. (the discussion regarding the expectation *i.e.*, $\theta = 0$, is analogous)

The switching curve is a direct application of monotonicity of $h_t^{\mathrm{P},\theta}(s,\cdot)$. Suppose there exists $\hat{s}^\dagger \in \mathcal{S}$ such that $h_t^{\mathrm{P},\theta}(s,\hat{s}^\dagger) > 0$ and $h_t^{\mathrm{P},\theta}(s,\hat{s}^\dagger - 1) \leq 0$. Then, we have $h_t^{\mathrm{P},\theta}(s,\hat{s}) > 0,\ \forall \hat{s} \geq \hat{s}^\dagger$ and $h_t^{\mathrm{P},\theta}(s,\hat{s}) \leq 0\ \forall \hat{s} < \hat{s}^\dagger$. We can let $\tau_t^\theta(s) = \hat{s}^\dagger - 1$. $\square$

A.2 Properties of entropic risk measure

Firstly, we show that the entropic risk measure satisfies the properties in Definition 3.1:

- (Risk-neutral) $\mu_0(\tilde{v}) = \mathbb{E}_\mathbb{P}[\tilde{v}]$.
- (Monotonicity) From Jensen's inequality, $[\mathbb{E}[\exp(\theta_1\tilde{v})]]^{\theta_2/\theta_1} \leq \mathbb{E}[\exp(\theta_2\tilde{v})]$, which implies that $[\mathbb{E}[\exp(\theta_1\tilde{v})]]^{1/\theta_1} \leq [\mathbb{E}[\exp(\theta_2\tilde{v})]]^{1/\theta_2}$ for any $\theta_1 \leq \theta_2$. Therefore, the first statement holds. Given the condition that $\theta > 0$, the function $\exp(\theta x)$ is increasing w.r.t. x. Then, according to the property of FSD, we have $\mathbb{E}[\exp(\theta\tilde{v}_1)] \geq \mathbb{E}[\exp(\theta\tilde{v}_2)]$, leading to the inference that $\mu_\theta(\tilde{v}_1) \geq \mu_\theta(\tilde{v}_2)$. In the case where $\theta = 0$, the inequality $\mathbb{E}[\tilde{v}_1] \geq \mathbb{E}[\tilde{v}_2]$ can be derived again, leveraging the property of FSD.
- (Normalization) $\mu_\theta(0) = \frac{1}{\theta}\log\mathbb{E}_\mathbb{P}\left[e^0\right] = 0$.
- (Translation invariance)

$$\mu_\theta(\tilde{v}+c) = \frac{1}{\theta}\log\mathbb{E}_\mathbb{P}\left[\exp\left(\theta\cdot(\tilde{v}+c)\right)\right] = \frac{1}{\theta}\log\left(\exp(\theta c)\,\mathbb{E}_\mathbb{P}\left[\exp\left(\theta\cdot\tilde{v}\right)\right]\right) = \mu_\theta(\tilde{v}) + c.$$

Properties 2 (Normalization) and 4 (Translation invariance) are easily derived according to the definition of $\mu_\theta(\tilde{v})$.

Moreover, we can prove that $(\theta,\tilde{v}) \mapsto \mu_{\frac{1}{\theta}}(\tilde{v})$ is jointly convex in $(\theta,\tilde{v})$:

- Using Hölder's inequality, we can show that $\mathbb{E}\left[\exp(\lambda\tilde{v}_1)\cdot\exp((1-\lambda)\tilde{v}_2)\right] \leq [\mathbb{E}[\exp(\tilde{v}_1)]]^\lambda \cdot [\mathbb{E}[\exp(\tilde{v}_2)]]^{1-\lambda}$. This implies that $log\,\mathbb{E}[\exp(\lambda\tilde{v}_1 + (1-\lambda)\tilde{v}_2)] \leq \lambda log\,\mathbb{E}[\exp(\tilde{v}_1)] + (1-\lambda) log\,\mathbb{E}[\exp(\tilde{v}_2)]$, *i.e.*, $log\,\mathbb{E}[\exp(\tilde{v})]$ is convex in $\tilde{v}$. The perspective of $log\,\mathbb{E}[\exp(\tilde{v})]$, that is $\mu_{\frac{1}{\theta}}(\tilde{v})$, preserves the convexity (the discussion of perspective operation can be found in §3.2.6 in Boyd and Vandenberghe (2004)).

A.3 Risk measure of terminal cost

In this section, we list the entropic risk measure of terminal costs under several different distributions. To approximate the value of definite integration, we use the Gaussian quadrature, where s_i is the i-th Gauss node in $[a,b]$ ($s_i = \frac{b-a}{2}\xi_i + \frac{a+b}{2}$) and w_i is the corresponding weight.

Table A.6: Entropic risk measure

Distribution, $\mathbb{P}^*$	$\mu_\theta(\tilde{v})$
Normal, $\mathcal{N}(\mu, \sigma^2)$	$\frac{1}{2}\sigma^2\theta$
Exponential, $exp(\lambda)$	$-\frac{1}{\theta}\log\left(1-\frac{\theta}{\lambda}\right)$, if $\theta < \lambda$
Uniform, U(a,b)	$\frac{1}{\theta}\log\left(\frac{1-exp(\theta(a-b))}{\theta(b-a)}\right)+b$
Truncated Lognormal (μ, θ^2) within $[a, b]$	$\frac{1}{\theta}log\left(\int_a^b \frac{1}{q_{Log}\sigma\sqrt{2\pi}x}\exp\left(\theta x - \frac{(\log\, x-\mu)^2}{2\sigma^2}\right)dx\right)$ $\approx \frac{1}{\theta}log\left(\frac{b-a}{2q_{Log}\sigma\sqrt{2\pi}}\sum_{i=1}^{n}\frac{w_i}{s_i}\exp\left(\theta s_i - \frac{(\log(s_i)-\mu)^2}{2\sigma^2}\right)\right)$, where $q_{Log} = \int_a^b \frac{1}{\sigma\sqrt{2\pi}x}\exp\left(-\frac{(\log\, x-\mu)^2}{2\sigma^2}\right)dx = \frac{b-a}{2\sigma\sqrt{2\pi}}\sum_{i=1}^{n}\frac{w_i}{s_i}\exp\left(-\frac{(\log(s_i)-\mu)^2}{2\sigma^2}\right)$
Empirical, $\hat{F}$	$\frac{1}{\theta}\log\left(\frac{1}{n}\sum_{i=1}^{n}\exp(\theta X_i)\right)$

A.4 Satisficing approach: bisection search for finding θ

Determining an appropriate risk preference parameter θ presents a significant practical challenge, as decision-makers often find it difficult to interpret and specify this parameter directly. To address this issue, we propose a strategy that aims to maximize risk tolerance while achieving a predefined satisfactory target, τ. This approach is aligned with the concept of robust satisficing, as discussed in Schwartz, et al. (2011) and D. Z. Long, et al. (2023). Setting a target or an acceptable level of sub-optimality is more intuitive and straightforward for decision-makers than specifying the complex parameter θ. In this context, we introduce the Satisficing optimal stopping problem (SOSP), which is formulated as follows:

$$\begin{aligned} \max \quad & \theta \\ \text{s.t.} \quad & \mu_\theta\left(\tilde{V}_0^{\boldsymbol{\pi}}\right) \leq \tau \\ & \theta \geq 0,\ \boldsymbol{\pi} \in \Pi_{\mathrm{MD}}. \end{aligned} \tag{A.28}$$

As established in Proposition 3.3, $\mu_\theta\left(\tilde{V}_0^{\boldsymbol{\pi}}\right)$ is monotonically increasing in θ. This property allows us to efficiently solve the satisficing problem using a bisection search algorithm (see Algorithm 1).

To facilitate practical implementation, we outline a step-by-step guideline:

- **Step 1:** Solve the risk-neutral version (NOSP), corresponding to $\theta = 0$. This yields the risk-neutral minimal cost, denoted as τ_0.

- **Step 2:** Using the benchmark level τ_0, select a satisfactory target $\tau = k \cdot \tau_0$, where $k \geq 1$. Note that the problem (A.28) is feasible if and only if $\tau \geq \tau_0$.

- **Step 3:** Employ a bisection search algorithm to determine the maximal value of θ that satisfies the constraint $\mu_\theta\left(\tilde{V}_0^{\boldsymbol{\pi}}\right) \leq \tau$.

This structured approach provides a practical and interpretable framework for decision-makers to balance risk tolerance and performance targets effectively.

Algorithm 1 Bisection Search θ for SOSP

Input Θ as the upper bound of θ, target value τ, and threshold ϵ.
Initialize $v_{low} = 0$.
Initialize $v_{high} = \Theta$.
while $v_{high} - v_{low} < \epsilon$ **do**
    $v_{mid} = (v_{low} + v_{high})/2$
    $u = \min_{\boldsymbol{\pi}} \mu_\theta\left(\tilde{V}^{\boldsymbol{\pi}}\right)$, *i.e.*, solve Problem (3.4) given $\theta = v_{mid}$.
    if $u \leq \tau$ **then**
        $v_{low} = v_{mid}$
    else
        $v_{high} = v_{mid}$
    end if
end while
Output: $\theta = v_{low}$

Following this, we utilize this satisficing approach to the extubation case with predictive information, as introduced in Section 3.5.1.2. The risk-neutral level is $\tau_0 = \mathbb{E}_{\mathbb{P}}[\tilde{V}^{\boldsymbol{\pi}^*(0)}] = 98.09$. The satisfactory target is chosen from $\{105, 110, 115, 120\}$. We utilize Monte Carlo simulations to generate $N = 10^6$ samples to evaluate the performance of NOSP-P and SOSP with predictive information (SOSP-P), denoted by $\{{}^jV\}$. Different from the main text, we evaluate the performance using two different metrics: the violation probability $\hat{\mathbb{P}}\left\{\tilde{V} > b\right\} = \frac{1}{N}\sum_{j=1}^N \mathbb{1}[{}^jV > b]$ and the conditional expected shortage $\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{V} - b \mid \tilde{V} > b\right] = \frac{1}{N\hat{\mathbb{P}}\{\tilde{V} > b\}}\sum_{j=1}^N \mathbb{1}[{}^jV > b] \cdot ({}^jV - b)$, with b selected from $\{105, 110, 115, 120, 125, 130, 135, 140\}$. The outcomes are detailed in Tables A.7 and A.8.

Table A.7: Violation probabilities $\hat{\mathbb{P}}(\tilde{V} > b)$ of NOSP-P and SOSP-P

	$\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{V}]$	$b = 105$	110	115	120	125	130	135	140
NOSP-P	98.09	0.43	0.38	0.34	0.30	0.26	0.23	0.20	0.17
SOSP-P									
$\tau = 105$	98.09	0.43	0.38	0.34	0.30	0.26	0.23	0.20	0.17
$\tau = 110$	100.07	0.44	0.39	0.34	0.30	0.26	0.22	0.19	0.16
$\tau = 115$	100.07	0.44	0.39	0.34	0.30	0.26	0.22	0.19	0.16
$\tau = 120$	100.70	0.45	0.40	0.35	0.30	0.26	0.22	0.19	0.16

A comparative analysis between NOSP-P and SOSP-P reveals that as the target τ escalates, SOSP-P achieves a more favorable risk profile, with a small loss in expected costs. SOSP-P strategically focuses on extreme values, leading to a slight enhancement in violation probability. In terms of conditional expected shortages, SOSP-P consistently outshines NOSP-P, with the magnitude of improvement intensifying in tandem with increases in τ. These findings reaffirm

Table A.8: Conditional expected shortages $\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{V}-b\middle|\tilde{V}>b\right]$ of NOSP-P and SOSP-P.

	$\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{V}]$	$b=105$	110	115	120	125	130	135	140
NOSP-P	98.09	35.17	33.95	32.84	31.83	30.93	30.09	29.32	28.61
SOSP-P									
$\tau=105$	98.09	35.12	33.91	32.80	31.83	30.94	30.14	29.38	28.68
$\tau=110$	100.07	31.35	29.98	28.69	27.51	26.43	25.44	24.56	23.74
$\tau=115$	100.07	31.41	30.02	28.75	27.59	26.53	25.53	24.61	23.75
$\tau=120$	100.70	30.93	29.52	28.26	27.10	26.03	25.06	24.15	23.34

the previously observed trade-off between expected performance and risk profile, analogous to ROSP-P.

A.5 Demographic and clinical variables of MIMIC-IV using for discharge decision making

We begin by selecting patient stays from MICU. The MIMIC-IV dataset contains 65,366 ICU admissions from 2008 to 2019, where 20,703 are MICU admissions. For each admission, we obtain demographic data such as age, gender, weight, race and admission time. During their hospitalization, we compile clinical data including those from physical condition monitors, such as heart rate, respiratory rate and blood pressure, and those of medication, such as Dopamine and Norepinephrine. Clinical variables were updated at intervals ranging from every 15 minutes to several hours, while results from labs were recorded at intervals from every six hours to nearly one day. The data preprocessing procedure involves several key steps:

1. **Variable Selection** Relevant clinical variables are selected based on their importance in monitoring patient conditions. These include vital signs (e.g., heart rate, blood pressure), lab results (e.g., hemoglobin, glucose), and ventilator parameters (e.g., tidal volume, PEEP level).

2. **Data Cleaning and Imputation** Missing data is handled through forward filling and linear interpolation. For variables with a high percentage of missing values, KNN imputation is applied after scaling the data. Abnormal values are filtered out using the interquartile range (IQR) method.

3. **Feature Engineering** New features are created, such as the Glasgow Coma Scale (GCS) score, which is derived from eye opening, verbal response, and motor response scores. Blood pressure values are consolidated from different sources (e.g., arterial and non-invasive measurements), and temperature values are standardized to Celsius.

4. **Vasopressor Information** Information on vasopressor use (e.g., Dopamine, Dobutamine, Norepinephrine, Epinephrine) is extracted and integrated into the dataset. This involves processing input events data to identify periods when these medications were administered.

5. **Time-Based Aggregation** Clinical data is aggregated over 12-hour time intervals, aligning with the decision frequency of our model. For the clinical variables, we calculate the average values within each time interval. For medication records, such as Dopamine, we calculate the maximum rate during the interval.

6. **Final Dataset Preparation** The processed data is saved into a final dataset. We also create features to denote whether patients are dead or readmitted within 30 days after discharge.

The coding and data analysis procedures are provided in the following GitHub repository: `https://github.com/LouisLou011/ROSP`.

After data preprocessing, we obtain 124,585 monitoring records that track the evolution of patients' status throughout their ICU stay. The definitions and descriptive statistics of the demographic and clinical variables are summarized in Table A.9. The last four rows correspond to Dopamine, Epinephrine, Norepinephrine, and Dobutamine, which are medications used to support cardiovascular function.

Table A.9: Definition and statistics of demographic and clinical variables from MIMIC-IV

Variables	Definition	Median (IQR)
Age	Patient's age at admission	62 (50 – 73)
Gender	Patient's gender	Male (55%)
Weight	Patient's weight (kg)	79.74 (66.0 – 97.2)
Heart Rate	Heart rate (beats per minute)	86.75 (75.33 – 98.62)
Respiratory Rate	Respiratory rate (breaths per minute)	20 (17.07 – 23.54)
Arterial O2 Pressure	Arterial oxygen pressure (mmHg)	105.33 (84.65 – 133.0)
Hemoglobin	Hemoglobin level (g/dL)	9.4 (8.2 – 10.9)
Arterial CO2 Pressure	Arterial carbon dioxide pressure (mmHg)	41.0 (36.8 – 46.8)
PH (Venous)	Venous blood pH	7.37 (7.33 – 7.41)
Hematocrit (serum)	Hematocrit level (%)	29.0 (25.6 – 33.5)
WBC	White blood cell count (10^3/µL)	10.5 (7.5 – 14.7)
Chloride (serum)	Serum chloride level (mmol/L)	103.0 (98.0 – 107.5)
Creatinine (serum)	Serum creatinine level (mg/dL)	1.1 (0.7 – 2.0)
Glucose (serum)	Serum glucose level (mg/dL)	126.0 (103.0 – 161.0)
Magnesium	Serum magnesium level (mg/dL)	2.1 (1.9 – 2.3)
Sodium (serum)	Serum sodium level (mmol/L)	139.0 (136.0 – 142.67)
PH (Arterial)	Arterial blood pH	7.4 (7.36 – 7.44)
Inspired O2 Fraction	Fraction of inspired oxygen (%)	46.38 (40.0 – 55.0)
BUN	Blood urea nitrogen (mg/dL)	26.3 (15.25 – 46.0)
Ionized Calcium	Ionized calcium level (mmol/L)	1.13 (1.09 – 1.17)
Total Bilirubin	Total bilirubin level (mg/dL)	0.64 (0.4 – 1.3)
Glucose (whole blood)	Whole blood glucose level (mg/dL)	133.7 (116.5 – 155.5)
Potassium (serum)	Serum potassium level (mmol/L)	4.0 (3.7 – 4.4)
HCO3 (serum)	Serum bicarbonate level (mmol/L)	24.0 (21.0 – 28.0)
Platelet Count	Platelet count (10^3/µL)	190.0 (120.0 – 272.0)
Prothrombin Time	Prothrombin time (seconds)	14.1 (12.6 – 16.9)
PTT	Partial thromboplastin time (seconds)	32.9 (28.4 – 42.9)
INR	International normalized ratio	1.3 (1.1 – 1.54)
Blood Pressure Mean	Mean blood pressure (mmHg)	77.09 (69.87 – 86.5)
Temperature	Body temperature (°C)	36.87 (36.61 – 37.19)
SaO2	Oxygen saturation (%)	96.75 (95.08 – 98.33)
GCS Score	Glasgow Coma Scale score	13.0 (8.67 – 15.0)
Dopamine	Dopamine use (10^{-2} µg/kg/min)	7.8 (0.6%)
Epinephrine	Epinephrine use (10^{-2} µg/kg/min)	0.2 (0.7%)
Norepinephrine	Norepinephrine use (10^{-2} µg/kg/min)	4.1 (18.7%)
Dobutamine	Dobutamine use (10^{-2} µg/kg/min)	2.5 (0.4%)

The last four rows correspond to medications administered to urgently support the cardiovascular system. We report the average of non-zero values (and the percentage of non-zero entries, respectively) instead.

SOFA score and transition matrix. We initiate the process by computing the SOFA score (ranging from 0 to 24). This score was developed by the Wroking Group on Sepsis-related Problems of the European Society of Intensive Care Medicine to assess organ dysfunction Vincent, et al. (1996). The distribution of patients' SOFA scores in all transition records is shown in Figure A.5. Due to the limited availability of data for higher scores, we merge all observations with SOFA scores greater than or equal to 16 into a single state. Thus, we define the following classes $\{0, 1, ..., 15, 16+, \text{Dealth}\}$, where "Death" is an absorbing state. The transition probability $P(i, j)$ from state i to state j can be estimated as $\frac{n_{ij}}{n_j}$, where n_i denotes the number of observations in state i and n_{ij} denotes the number of transitions from state i to state j. To ensure data stability, we exclude admissions whose LOS is below 5th percentile, as a short ICU hospitalization often indicates unstable status and may introduce noise and bias into the analysis. The estimated transition probabilities are available on `https://github.com/LouisLou011/ROSP`.

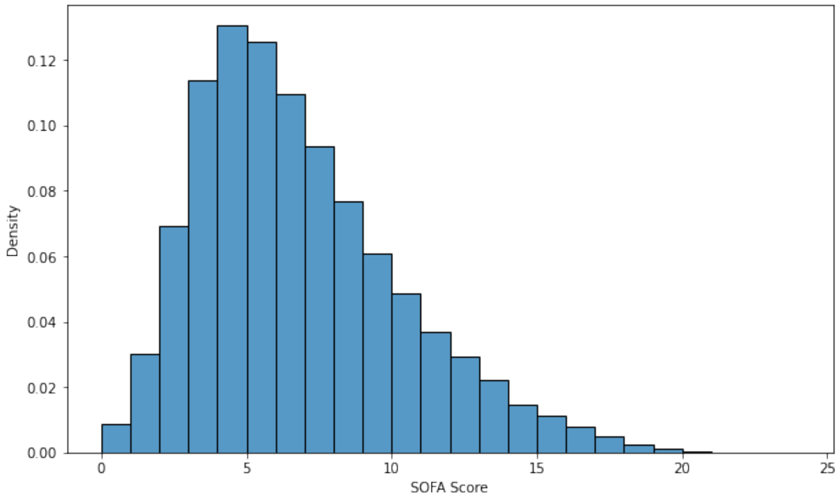


Figure A.5: Distribution of SOFA score in MICU using MIMIC-IV

Terminal cost. The terminal cost is composed of two components: 30-day out-of-ICU mortality and readmission cost. To estimate these, we employ logistic regression to predict 30-day out-of-ICU mortality and readmission rates. Below, we outline the detailed procedure:

1. We begin by selecting patients who are admitted to MICU for the first time, excluding any readmission cases. We further exlude patients who experience in-ICU mortality.

2. The dependent variable is defined as whether a patient is readmitted to the ICU or dies within 30 days after their initial discharge. The explanatory variables consist of the latest recorded clinical and demographic measurements prior to the patient's discharge.

3. The dataset is randomly partitioned into 80% of training data and 20% of testing data.

4. Two logistic regression models are trained on the training dataset to predict 30-day out-of-ICU mortality and admission. The model is specified as follows:

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 \mathrm{SOFA} + \boldsymbol{\beta}_2^\top \boldsymbol{x},$$

   where p is either the mortality rate or the readmission rate. Notice that these events are not mutually exclusive, as a patient may die during readmission. $\boldsymbol{x}$ includes other clinical and demographic variables.

5. The performance of these classification model is evaluated on the testing data, and the results are presented in Table A.10.

6. Then, we estimate the parameters for the terminal cost. Given that there are very few samples discharged in severe status, such as SOFA score 16+, we further group patients into 4 groups: SOFA scores $\{0, 1\}$, $\{2, 3\}$, $\{4, 5, 6, 7\}$, $\{8, ..., 24\}$. For each group, we calculate

the average value of explanatory variables within each group, denoted as $\bar{\boldsymbol{x}}_{\{0,1\}}$, $\bar{\boldsymbol{x}}_{\{2,3\}}$, $\bar{\boldsymbol{x}}_{\{4,5,6,7\}}$, and $\bar{\boldsymbol{x}}_{\{8,\ldots,24\}}$.

7. Using the predicted parameters $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\boldsymbol{\beta}}_2$, we can estimate the out-of-ICU motarlity or readmission rate of each state. For example, the mortality or readmission rate for SOFA score 1 is calculated as:

$$\log \frac{p_1}{1-p_1} = \beta_0 + \beta_1 \cdot 1 + \boldsymbol{\beta}_2^\top \bar{\boldsymbol{x}}_{\{0,1\}}.$$

Following these procedures, we estimate the 30-day out-of-ICU mortality and readmission rates for each SOFA score group. Table A.11 presents estimated rates, which serve as inputs for the numerical case discussed in Section 3.5.2.

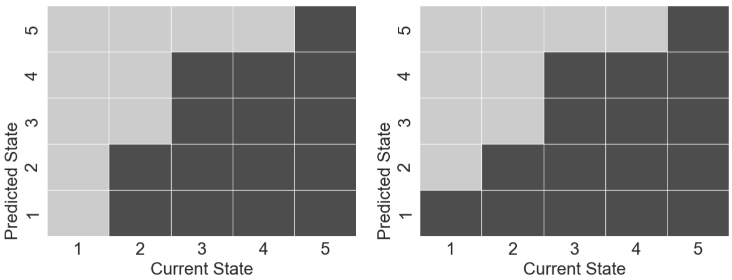
Table A.10: Prediction performance of 30-day out-of-ICU mortality and readmission

	ROC-AUC		Accuracy		F1 score	
	Training	Testing	Training	Testing	Training	Testing
Mortality	82%	83%	87%	89%	87%	88%
Readmission	63%	64%	77%	74%	79%	78%

Table A.11: Estimated 30-day out-of-ICU mortality and readmission

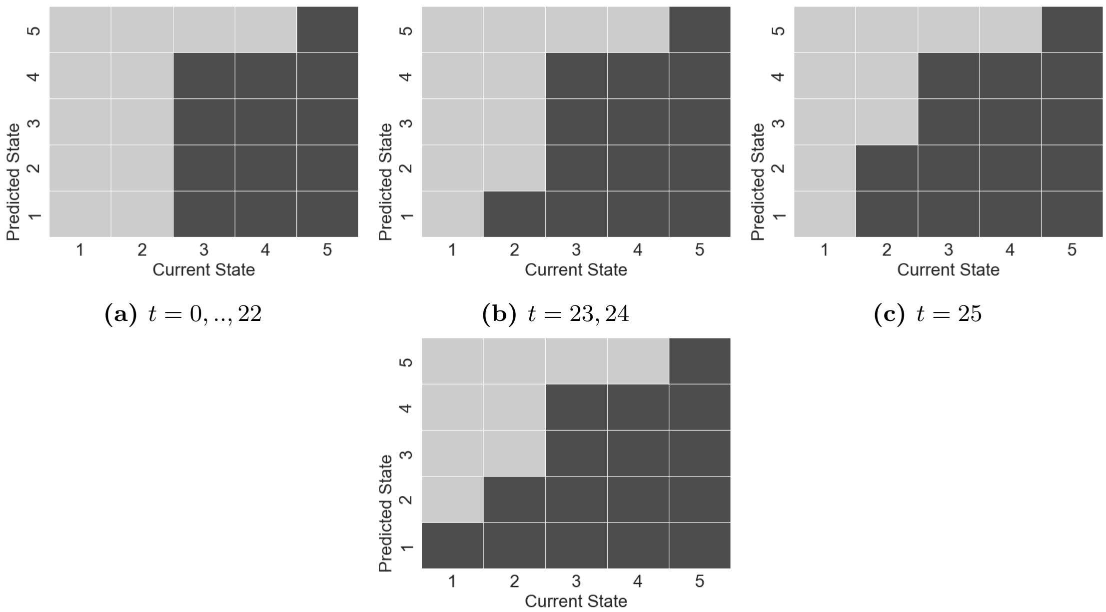
SOFA score	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16+
mortality(%)	1.66	2.30	3.19	4.40	6.05	8.27	11.19	14.99	19.78	25.64	32.54	40.28	48.55	56.89	64.86	72.08	93.25
readmission(%)	10.20	11.02	11.89	12.82	13.81	14.86	15.98	17.17	18.42	19.75	21.14	22.61	24.14	25.75	27.43	29.17	38.74

A.6 Additional figures



(a) $t = 0, .., 25$ **(b)** $t = 26$

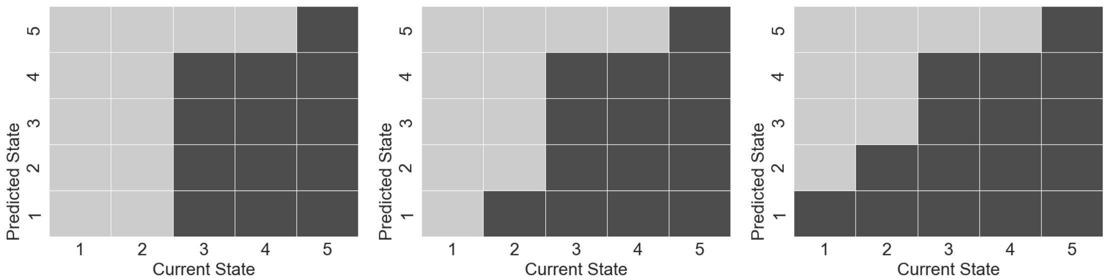
Figure A.6: Optimal policy of ROSP-P given $\theta = 0.005$. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.



(a) $t = 0, .., 22$ **(b)** $t = 23, 24$ **(c)** $t = 25$

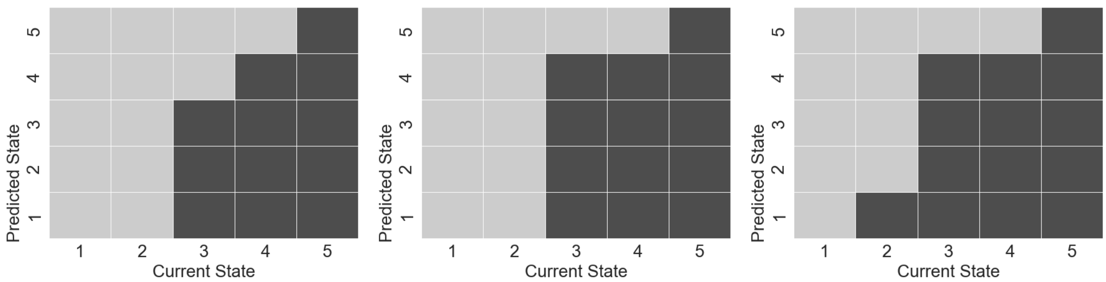
(d) $t = 26$

Figure A.7: Optimal policy of ROSP-P given $\theta = 0.01$. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.

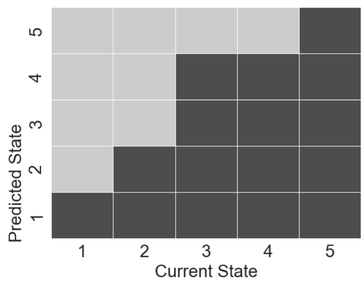


(a) $t = 0, .., 23$ **(b)** $t = 24, 25$ **(c)** $t = 26$

Figure A.8: Optimal policy of ROSP-P given $\theta = 0.015$. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.

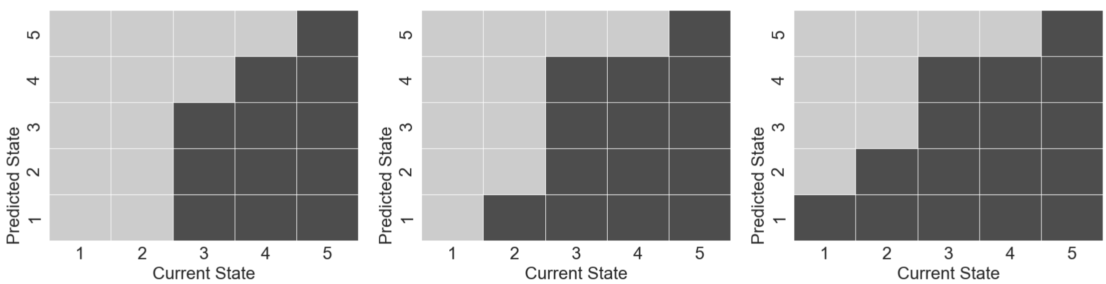


(a) $t = 0, .., 22$ **(b)** $t = 23$ **(c)** $t = 24, 25$



(d) $t = 26$

Figure A.9: Optimal policy of ROSP-P given $\theta = 0.025$. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.



(a) $t = 0, .., 23$ **(b)** $t = 24, 25$ **(c)** $t = 26$

Figure A.10: Optimal policy of ROSP-P given $\theta = 0.03$. The dark gray grid indicates extubation, *i.e.*, $a = 1$, while the light gray denotes continuation, *i.e.*, $a = 0$.

4 Robust epidemiological analytics

Building on the risk-aware decision frameworks developed in Chapter 3, we now shift our focus from individual-level clinical decisions to a broader public health context—specifically, managing the spread of infectious diseases through robust and adaptive intervention strategies (ESCAP, et al., 2020). While the previous chapter addressed optimal stopping decisions in critical care settings, the present chapter considers population-level interventions under epidemic uncertainty. Here, the central challenge is not determining when to stop treatment for a patient, but rather how to coordinate interventions across an entire community in a timely and robust manner.

Unlike the finite-horizon MDP models discussed previously, epidemic dynamics are better represented as queueing systems in which populations evolve across health states (e.g., susceptible, exposed, infected, recovered). The complexity of this setting is heightened by shared resource constraints and interdependencies across compartments—particularly when capacity limitations must be managed across heterogeneous population groups. Inadequate planning or delayed responses have repeatedly resulted in resource shortages—especially for ICU beds, ventilators, and medical staff—which in turn have amplified the morbidity and mortality associated with the virus.

Designing prescriptive analytics models to inform government policies in such contexts presents a dual challenge. First, effective intervention planning requires accurate short- to medium-term forecasts of the disease trajectory, which demands stochastic models capable of capturing the uncertainty in epidemic progression. Second, integrating such forecasts into an optimization model that remains computationally tractable is nontrivial, particularly when the dynamics involve recursive, decision-dependent transitions.

Although deterministic compartmental models remain widely used due to their simplicity and ease of interpretation, they fall short in capturing the nuances of stochastic environments—such as random patient inflow and fluctuating healthcare demand. These shortcomings are especially apparent in scenarios involving limited medical resources. For instance, when patient arrivals and discharges are volatile, the availability of critical care beds becomes a bottleneck, often resulting in treatment delays and prolonged queues (see, for example, Xie, et al., 2023).

To effectively manage such uncertainty, robust modeling approaches are essential—both on the forecasting side and within the optimization framework. Yet real-world deployment of these approaches remains difficult. Accurate long-term prediction is inherently challenging due to latent factors such as viral mutations, behavioral shifts, and reporting delays. Furthermore, capturing high-dimensional dependencies across multiple compartments exacerbates model complexity. These difficulties are further compounded by the recursive structure of compartmental dynamics, in which transition rates depend intricately on both decisions and current state variables.

This work contributes towards addressing these issues. We present a novel robust epidemiological prescriptive analytics framework, leveraging stochastic compartmental dynamics to account for uncertainty. Our proposed model aims to minimize the collective *ambiguity tolerances*, quantifying the violation of a set of healthcare-related targets such as, *inter alia*, the number of peak infections and deaths, hospital bed capacities, and total implementation costs. The objective of minimizing the ambiguity tolerance is related to models that minimize variants of the riskiness index (*e.g.,* Aumann and Serrano, 2008; D. B. Brown and Sim, 2009; Zhou, et al., 2022a). The methodology has been successfully implemented in several applications in the context of healthcare resource management (Xie, et al., 2023; Zhou, et al., 2022b). The decision criterion we adopt is also closely related to robust satisficing models (D. Z. Long, et al., 2023). Compared with the classical robust optimization approaches (Bertsimas and Sim, 2004; Ben-Tal, et al., 2010), the satisficing principle does not require a specific ambiguity set. From the satisficing perspective, our model aims to achieve the epidemiological performance targets as well as possible under all possible distributions, which can be associated with having the lowest tolerance for violating targets under ambiguity. Our robust epidemiological optimization model tractably incorporates forecasts from the stochastic dynamic of the pandemics. It can be reformulated to a deterministic non-convex optimization problem because the certainty equivalent of the random compartment populations can be calculated recursively. The asymptotic Gaussian property enables us to apply a mild Gaussian approximation of the certainty equivalent, thereby further enhancing the computational tractability of the robust model. It is worth noticing that the proposed robust epidemiological optimization model can be easily modified to a cost-driven model by fixing the ambiguity tolerance variable and minimizing a cost function, and the proposed solution approach can be adopted. Our proposed framework is not limited to a specific type of compartmental model. It accommodates a versatile network structure, empowering policymakers to tailor their approach based on local epidemic dynamics and preventive measures.

The contributions of our work are summarized as follows. [1]

[1]This chapter draws on content from the pre-print article: Fu, C., Zhou, M., Lou, Z., Xie, J., Sim, M., and Tan, K.B. (2024). Analytics with Robust Epidemiological Compartmental Optimization Models. Available at: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3869521`.

Authorship and contribution details for the pre-print (relevant to this chapter's content): Chenyi Fu, Minglong Zhou, and Zhiyuan Lou (PhD candidate, this dissertation) made equal contributions to the article's writing and theoretical development of the robust epidemiological models. Additionally, Chenyi Fu took primary responsibility for the pre-print's numerical work. Note that the numerical results using Singapore data (included in the pre-print) are not displayed in this chapter, consistent with the scope of the dissertation. This work has not been incorporated into any other dissertation before.

- **Incorporate stochastic epidemiological dynamics into optimization.** We propose a general stochastic compartmental model that incorporates decision-dependent compartments such as the tested and vaccinated populations. The related decisions directly influence the progression of epidemiology, subsequently affecting the demand for public health services.

- **Assess solution robustness with an ambiguity tolerance metric.** We introduce an ambiguity tolerance measure and propose a robust optimization model that minimizes the collective ambiguity tolerance. Both the ambiguity tolerance measure and the robust model are novel in epidemiological analytics.

- **Tractable approximation of the robust compartmental optimization model.** We demonstrate that any subset of compartments is asymptotically Gaussian, lending support to a Gaussian approximation of the certainty equivalent. This substantially reduce the computational complexity.

- **Outperforms the deterministic model in a real-world application.** We conduct a case study based on real-world data from Singapore. Our numerical results highlight the efficacy of the proposed robust epidemiological modeling framework. Notably, our approach outperforms the deterministic model, leading to a reduction in both hospitalized cases and deaths, as well as low probabilities of constraint violations.

The remainder of this chapter is organized as follows. In Section 4.1, we present the deterministic epidemiological dynamics and deterministic optimization model. In Section 4.2, we present the stochastic epidemiological model and present a robust optimization model based on an ambiguity tolerance measure. In Section 4.3, we detail the reformulation of the robust model and suitable approximations. We also use a real-world dataset to test our model; for the detailed result, please refer to our pre-print version (Fu, et al., 2022). In Section 4.4, we conclude this chapter. Proofs are relegated to Appendix B.1.

4.1 Deterministic compartmental optimization model

In this section, we formulate a general deterministic compartmental optimization model. We focus on a general model with K population compartments represented by the set of labels, $\mathcal{K}$, and the population is segmented into J different groups, *e.g.,* based on age, health status, spatial location, or other characteristics. The population size of group j is denoted by N_j and the total population size over all groups is denoted by N.

4.1.1 Deterministic compartmental model

The deterministic compartmental optimization model formulates the dynamics of the populations of all compartments and groups over the time horizon of length T based on the system's initial state at time $t = 1$, *e.g.,* given model parameters and initial populations. Let $\bar{Z}^n_{j,t}$ be the group j population in compartment n at time t, for all $n \in \mathcal{K}$, $j \in [J]$, and $t \in [T]$. Governments and healthcare practitioners can employ non-pharmacological/pharmacological intervention policies to control the flow of individuals among populations to some extent (see Example 4.1). For all $n, m \in \mathcal{K}$, $j \in [J]$, and $t \in [T]$, we use $x^{m,n}_{j,t}$ to denote a decision variable representing the flow of group j population from compartment m to n at period t, which can be a result of non-pharmacological/pharmacological intervention policies (*e.g.*, vaccination and quarantine). In our model, group j is defined by patient characteristics (*e.g.*,age), and we assume no transitions between groups over the planning horizon.

Example 4.1 (SVEIR Model)**.** *The SVEIR model considers five population compartments, labeled by:* Susceptible*,* Vaccinated*,* Exposed*,* Infectious*, and* Removed*, i.e.,* $\mathcal{K} = \{S, V, E, I, R\}$*. Let* $\bar{\boldsymbol{Z}}^S, \bar{\boldsymbol{Z}}^V, \bar{\boldsymbol{Z}}^E, \bar{\boldsymbol{Z}}^I, \bar{\boldsymbol{Z}}^R$ *be the corresponding compartment variables representing the sizes of these compartments. The* Removed *compartment represents the deaths and recovered individuals that become immune to the disease, at least for a substantial amount of time greater than* T*. To model the impact of vaccination, we include the decision variables* $x^{S,V}_{j,t} \geq 0$*, for* $j \in [J]$ *and* $t \in [T]$*, to denote the number of scheduled vaccinations for the* j*th group at period* t*. The vaccinated population would be less likely to be exposed in subsequent periods.*

The populations in different compartments in groups $j \in [J]$ *at time* $t \in [T-1]$ *are as follows:*

$$\bar{Z}^n_{j,t+1} = \sum_{m \in \mathcal{K}/\{n\}} x^{m,n}_{j,t} + \sum_{m \in \mathcal{K}} \hat{q}^{m,n}_{j,t} \left(\bar{Z}^m_{j,t} - \sum_{m' \in \mathcal{K}/\{m\}} x^{m,m'}_{j,t} \right), \quad \forall\, n \in \mathcal{K}, j \in [J], t \in [T-1], \tag{4.1}$$

where $\hat{q}^{m,n}_{j,t}$ *is the transition probability from compartment* m *to* n *for group* j *at time* t*, which satisfies* $\sum_{n \in \mathcal{K}} \hat{q}^{m,n}_{j,t} = 1$*.*

The transition rates between compartments must be estimated or learned from historical information and expert opinions. These transition rates are often endogenous because they could depend on interactions among groups and compartments. For example, in the above SVEIR model, the susceptible cases may be exposed after interactions with the infected cases. Hence, the transition rate, $\hat{q}^{S,E}_{j,t}$, is an affine function of the number of infected cases and how susceptible individuals interact with them, and it is often modeled as

$$\hat{q}_{j,t}^{S,E} := \sum_{k\in[J]} \hat{\zeta}_{k,j} \bar{Z}_{k,t}^{I}.$$

Specifically, the parameter $\hat{\zeta}_{k,j}$ relates to the estimated interaction-transmission rate between groups k and j. For any group $k \in [J]$, its interaction-transmission rate can be expressed as $\hat{\zeta}_{k,j} = \frac{\hat{\beta}_k \hat{w}_{k,j}}{N_k}$, where $\hat{\beta}_k$ is the transmission rate, and $\hat{w}_{k,j}$ is the interaction rate with the jth group. Then, $\sum_{k\in[J]} \hat{\zeta}_{k,j} \bar{Z}_{k,t}^{I}$ represents the overall estimated interaction-transmission rate of group $j \in [J]$. Different from the conventional compartment model, Equation (4.1) presents a general formulation of epidemiological dynamics including decisions, which can formulate different compartmental model structures. In the following example, we illustrate that one can use Equation (4.1) to formulate the Delphi model presented by M. L. Li, et al. (2023).

Example 4.2 (Delphi Model)**.** *The Delphi model contains 11 compartments and forecasts the numbers of infected individuals and deaths in a granular fashion (see Figure 4.1). The compartments are labeled as the susceptible ($\bar{\boldsymbol{Z}}^S$), exposed ($\bar{\boldsymbol{Z}}^E$), infected ($\bar{\boldsymbol{Z}}^I$), undetected recovered ($\bar{\boldsymbol{Z}}^{UR}$), detected and hospitalized recovered ($\bar{\boldsymbol{Z}}^{DHR}$), detected and quarantined recovered ($\bar{\boldsymbol{Z}}^{DQR}$), undetected death ($\bar{\boldsymbol{Z}}^{UD}$), detected and hospitalized death ($\bar{\boldsymbol{Z}}^{DHD}$), detected and quarantined death ($\bar{\boldsymbol{Z}}^{DQD}$), recovered ($\bar{\boldsymbol{Z}}^R$), and deceased ($\bar{\boldsymbol{Z}}^D$). The discrete Delphi model can be written in the form of* (4.1)*, where the compartment transition probabilities are defined accordingly. For example,*

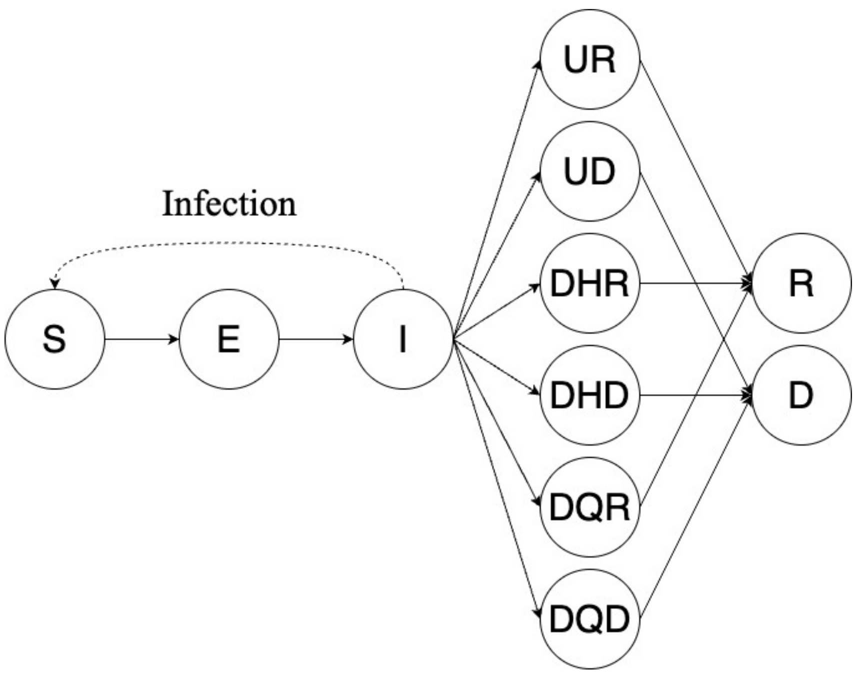


Figure 4.1: Framework of the Delphi model.

$$
\begin{aligned}
\bar{Z}^{S}_{j,t+1} &= (1 - \hat{q}^{S,E}_{j,t})\bar{Z}^{S}_{j,t} = \hat{q}^{S,S}_{j,t}\bar{Z}^{S}_{j,t} \\
\bar{Z}^{E}_{j,t+1} &= (1 - \hat{q}^{E,I}_{j,t})\bar{Z}^{E}_{j,t} + \hat{q}^{S,E}_{j,t}\bar{Z}^{S}_{j,t} = \hat{q}^{E,E}_{j,t}\bar{Z}^{E}_{j,t} + \hat{q}^{S,E}_{j,t}\bar{Z}^{S}_{j,t},
\end{aligned}
$$

for some overall transition rate $\hat{q}^{S,E}_{j,t}$ *and incubation rate* $(1 - \hat{q}^{E,I}_{j,t})$.

Bertsimas, et al. (2022) further integrate vaccine distribution into the above model and present a Delphi-V model. This is achieved by incorporating additional compartments and compartment variables $\bar{\boldsymbol{Z}}^{SV}$, $\bar{\boldsymbol{Z}}^{EV}$, $\bar{\boldsymbol{Z}}^{IV}$, *and* $\bar{\boldsymbol{Z}}^{RV}$ *representing the numbers of susceptible and vaccinated cases, exposed and vaccinated cases, infected and vaccinated cases, and immune cases. The vaccinated cases have a lower transmission rate, affecting the rest of the dynamics. For example, the dynamics associated with the first and the second compartments become*

$$
\begin{aligned}
\bar{Z}^{S}_{j,t+1} &= (1 - \hat{q}^{S,E}_{j,t})(\bar{Z}^{S}_{j,t} - x^{S,SV}_{j,t}) = \hat{q}^{S,S}_{j,t}(\bar{Z}^{S}_{j,t} - x^{S,SV}_{j,t}) \\
\bar{Z}^{E}_{j,t+1} &= (1 - \hat{q}^{E,I}_{j,t})\bar{Z}^{E}_{j,t} + \hat{q}^{S,E}_{j,t}(\bar{Z}^{S}_{j,t} - x^{S,SV}_{j,t}) = \hat{q}^{E,E}_{j,t}\bar{Z}^{E}_{j,t} + \hat{q}^{S,E}_{j,t}(\bar{Z}^{S}_{j,t} - x^{S,SV}_{j,t}),
\end{aligned}
$$

where $x^{S,SV}_{j,t}$ *is the number of vaccines distributed into group* j *at time* t.

4.1.2 Optimizing decision using deterministic compartmental model

Embedding the epidemiological dynamics in an optimization problem can help anticipate how the non-pharmacological/pharmacological decisions would affect the disease propagation. Besides the flow decisions $\boldsymbol{x}$, we include some additional decisions $\boldsymbol{y} \in \mathbb{R}^{N_y}$, that could affect the flow decisions indirectly (*e.g.,* vaccine center facility location decisions, hospital capacity planning decisions and others). We use the set $\bar{\mathcal{X}}$ to represent a convex feasibility set of $(\boldsymbol{x}, \boldsymbol{y})$. Specifically, we define

$$
\bar{\mathcal{X}} := \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{K^2JT}_{+} \times \mathbb{R}^{N_y} \;\middle|\; \begin{array}{l} g_i(\boldsymbol{x}, \boldsymbol{y}) \le 0 \quad \forall i \in [I_g] \\ \boldsymbol{x} \in \mathcal{X}, \boldsymbol{y} \in \mathcal{Y} \end{array} \right\},
$$

where $\mathcal{X}$ and $\mathcal{Y}$ contain only mixed-integer linear constraints, and g_i is a convex function for any $i \in [I_g]$ (I_g is the number of policy constraints specified by policymakers). The interaction-transmission rate, $\hat{q}^{m,n}_{j,t}$, is modeled as an affine function of $\bar{\mathbf{Z}}$ and $\boldsymbol{x}$, denoted by $\eta^{m,n}_{j,t}$. Set $\mathcal{D}$ records all pairs of indices of two compartments where the transition rates between these compartments are decision-dependent. For $(m, n) \notin \mathcal{D}$, the transition rate $\hat{q}^{m,n}_{j,t}$ is a fixed constant. We now present the baseline deterministic optimization model, which is embedded with the compartment dynamics described in (4.1):

$$
\begin{aligned}
Z_0 = \min \quad & \sum_{n \in \mathcal{K}} \mathbf{C}^n \bullet \bar{\mathbf{Z}}^n + \boldsymbol{c}_1^\top \boldsymbol{x} + \boldsymbol{c}_2^\top \boldsymbol{y} \\
\text{s.t.} \quad & \text{Dynamics constraint (4.1)} \\
& f_i\left(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}\right) \le 0 && \forall i \in [I_f] \\
& \hat{q}_{j,t}^{m,n} = \eta_{j,t}^{m,n}(\bar{\mathbf{Z}}, \boldsymbol{x}) && \forall (m,n) \in \mathcal{D}, j \in [J], t \in [T] \\
& \bar{\mathbf{Z}}^n \in \mathbb{R}_+^{J \times T} && \forall n \in \mathcal{K} \\
& (\boldsymbol{x}, \boldsymbol{y}) \in \bar{\mathcal{X}}.
\end{aligned}
\tag{4.2}
$$

The objective function minimizes the total operational cost, where $\mathbf{C}^n$, $\boldsymbol{c}_1$, and $\boldsymbol{c}_2$ are the unit cost for compartment n, flow decision $\boldsymbol{x}$, and additional decision $\boldsymbol{y}$, respectively. According to the policymakers' requirements, the objective function can include, for example, costs of healthcare resources and economic losses due to infections and deaths. $f_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) \le 0$ dictates the operational constraints associated with decision variables and (forecast) compartment sizes, for any $i \in [I_f]$ (I_f is the number of operational constraints). For tractability, we restrict ourselves to only consider the case where f_i is a linear function in the following form.

$$
f_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) = \sum_{n \in \mathcal{K}} \mathbf{A}_i^n \bullet \bar{\mathbf{Z}}^n + \boldsymbol{b}_{i,1}^\top \boldsymbol{x} + \boldsymbol{b}_{i,2}^\top \boldsymbol{y} - h_i \qquad \forall i \in [I_f],
\tag{4.3}
$$

1. **Decision budget constraints.** We can use linear constraints to enforce operational budgets for certain intervention decisions. In general, we can have the following budget constraint,

$$
f_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) = \sum_{m \in \mathcal{K}_1} \sum_{n \in \mathcal{K}_2} \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} x_{j,t}^{m,n} - h_{\mathcal{J},\mathcal{T}}^{\mathcal{K}_1,\mathcal{K}_2},
$$

   where $\mathcal{K}_1$ and $\mathcal{K}_2$ are given subsets of $\mathcal{K}$, $\mathcal{J}$ is a given subset of $[J]$, and $\mathcal{T}$ is a given subset of $[T]$. $h_{\mathcal{J},\mathcal{T}}^{\mathcal{K}_1,\mathcal{K}_2}$ is the associated budget. For example, in the SVEIR model, variable $x_{j,t}^{S,V}$ is the number of vaccines allocated to group j at time t. The total vaccine distribution at time t should be no greater than the total number of vaccines available at time t (denoted as $h_t^{S,V}$). The number of vaccines allocated to group j population should be no greater than the number of susceptible cases. These can translate to the following constraints,

$$
f_i(\mathbf{Z}, \boldsymbol{x}, \boldsymbol{y}) = \sum_{j \in [J]} x_{j,t}^{S,V} - h_t^{S,V}.
$$

2. **Compartment capacity constraints.** Similar to decision variables, some compartments may have capacity limits. For example, a compartment representing the number of hos-

pitalized cases should be limited by the number of available beds. We can allow capacity restriction constraints in the following form:

$$f_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) = \sum_{n\in\mathcal{K}_1}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}} \bar{Z}^n_{j,t} - h^{\mathcal{K}_1}_{\mathcal{J},\mathcal{T}}$$

for some $\mathcal{K}_1 \subseteq \mathcal{K}$, $\mathcal{J} \subseteq [J]$, and $\mathcal{T} \subseteq [T]$ and $h^{\mathcal{K}_1}_{\mathcal{J},\mathcal{T}}$ is the capacity for compartment set $\mathcal{K}_1$.

3. **Fairness.** Some decisions, such as vaccine distribution, test kit allocation, and ward allocation, may be subject to fairness concerns among population groups. Some examples of fairness concerns in vaccine allocation are considered in Bertsimas, et al. (2022). We can incorporate similar fairness constraints as (mixed-integer) linear constraints. For example, the following fairness constraint ensures that the number of resources distributed to group j does not deviate too much from that to group k,

$$\begin{aligned}\bar{f}_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) &= \sum_{m\in\mathcal{K}_1}\sum_{n\in\mathcal{K}_2}\sum_{t\in\mathcal{T}}\left(\frac{x^{m,n}_{j,t}}{N_j} - \frac{x^{m,n}_{k,t}}{N_k}\right) - \bar{h}_{j,k} \quad \forall j,k\in[J]\\ \underline{f}_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) &= \sum_{m\in\mathcal{K}_1}\sum_{n\in\mathcal{K}_2}\sum_{t\in\mathcal{T}}\left(\frac{x^{m,n}_{k,t}}{N_k} - \frac{x^{m,n}_{j,t}}{N_j}\right) + \underline{h}_{j,k} \quad \forall j,k\in[J]\end{aligned}$$

for some $\mathcal{K}_1 \subseteq \mathcal{K}$, $\mathcal{K}_2 \subseteq \mathcal{K}$, and $\mathcal{T} \subseteq [T]$. The hyperparameters $\bar{h}_{j,k} \in [0,1]$ and $\underline{h}_{j,k} \in [-1,0]$ can be set by policy-makers, whose values control the trade-off between efficiency and fairness, *i.e.*, when $\bar{h}_{j,k} - \underline{h}_{j,k}$ tends to zero (two), the proportion of resources distributed to different groups should be the same (different) reflecting fairness (efficiency).

The computational tractability of Problem (4.2) heavily depends on the form of the transition rate function. In the third constraint of (4.2), we focus on affine transition rates, *i.e.*, a transition rate is an affine function of $\bar{\mathbf{Z}}$ and $\boldsymbol{x}$,

$$\eta^{m,n}_{j,t}(\bar{\mathbf{Z}}, \boldsymbol{x}) = \sum_{m'\in\mathcal{K}}\sum_{k\in[J]}\left(\beta^{m,n,m'}_{k,j,t}\bar{Z}^{m'}_{k,t} + \gamma^{m,n,m'}_{k,j,t}\sum_{n'\in\mathcal{K}} x^{m',n'}_{k,t}\right) + \alpha^{m,n}_{j,t}, \tag{4.4}$$

for some given parameters $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\alpha}$. The transition rate function is decision-dependent and would incur bilinear terms when plugged into the epidemiological dynamics (4.1). For example, in the SVEIR model, the transmission rate from the susceptible ($\bar{\mathbf{Z}}^S$) to the exposed ($\bar{\mathbf{Z}}^E$) depends on the number of infections ($\bar{\mathbf{Z}}^I$), *i.e.,*

$$\eta^{S,E}_{j,t}(\bar{\mathbf{Z}}, \bar{\boldsymbol{x}}) = \sum_{k\in[J]}\beta^{S,E,I}_{k,j,t}\bar{Z}^I_{k,t}.$$

However, even when $\eta_{j,t}^{m,n}$ is an affine function, addressing Problem (4.2) remains a formidable task, as pointed out by Bertsimas, et al. (2022). That is, we would encounter bilinear terms when evaluating the number of transitions from the susceptible to the exposed compartment, $\hat{q}_{j,t}^{S,E}\bar{Z}_{j,t}^{S} = \bar{Z}_{j,t}^{S}\sum_{k\in[J]}\beta_{k,j,t}^{S,E,I}\bar{Z}_{k,t}^{I}$. As a result, Problem (4.2) poses as a large-scale, non-convex optimization problem that is challenging to solve optimally within polynomial time. The simulation-optimization iterative approach can be employed (refer to, for instance, Bertsimas, et al. (2022)) to create a nearly optimal solution, the details of which are presented in the appendix of Fu, et al. (2022). The subsequent example offers a demonstrative case for implementing our general model.

Example 4.3 (Vaccine centers facility location with SVEIR dynamics)**.** *As a simple illustration of the epidemiological optimization model* (4.2)*, we extend the SVEIR model in Example 1 and include vaccine center facility locations. A group signifies a specific region (such as a state, city, or district). In this context, the model evolves into a spatial compartment model. The planner minimizes the total cost of vaccine allocation and running vaccine centers. Let* $\boldsymbol{y} = (\boldsymbol{y}^{f\top}, \boldsymbol{y}^{c\top}, \boldsymbol{y}^{s\top})^\top$ *be the additional decisions relating to the vaccine centers, where* $\boldsymbol{y}^f \in \{0,1\}^J$, $\boldsymbol{y}^c \in \mathbb{Z}_+^{JT}$, *and* $\boldsymbol{y}^s \in \mathbb{R}_+^{J^2T}$. *The binary decision variable* y_j^f *equals one if the vaccine center located in region* j *is open. The decision variable* $y_{j,t}^c$ *is the number of vaccines ordered by the vaccine center in region* j *at time* t, *and* $y_{j,k,t}^s$ *is the number of vaccines from the vaccine center in region* j *used to satisfy the vaccine requirement by region* k *at time* t. *In the SVEIR setup, the flow decision* $x_{j,t}^{S,V}$ *is the total number of vaccines allocated to the susceptible population in region* j *at time* t. *The deterministic epidemiological optimization model can be written as*

$$
\begin{aligned}
Z_0 = \min \quad & \boldsymbol{c}_1^\top \boldsymbol{x} + \boldsymbol{c}_f^\top \boldsymbol{y}^f + \boldsymbol{c}_c^\top \boldsymbol{y}^c + \boldsymbol{c}_s^\top \boldsymbol{y}^s \\
\text{s.t.} \quad & \textit{Dynamics constraint } (4.1) \\
& \sum_{j\in[J]} \bar{Z}_{j,t}^{I} - h_t \leq 0 && \forall t \in [T] \\
& \hat{q}_{j,t}^{S,E} = \eta_{j,t}^{S,E}(\bar{\mathbf{Z}}, \boldsymbol{x}) && \forall j \in [J], t \in [T] \\
& \bar{\mathbf{Z}}^n \in \mathbb{R}_+^{J\times T} && \forall n \in \mathcal{K} \\
& (\boldsymbol{x}, \boldsymbol{y}) \in \bar{\mathcal{X}}.
\end{aligned}
$$

The four cost parameters in the objective function represent the variable cost of vaccination, the fixed cost of facility location, the variable cost of the vaccine stocks, and the transportation cost from region j *to* k. *The second constraint is a linear constraint ensuring that the number of* *active infections is under a target threshold. The feasibility set $\bar{\mathcal{X}}$ is given by*

$$\bar{\mathcal{X}} = \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}_+^{K^2JT} \times \mathbb{R}^{J+JT+J^2T} \;\middle|\; \begin{array}{ll} x_{k,t}^{S,V} - \sum\limits_{j\in[J]} y_{j,k,t}^s \leq 0 & \forall k \in \mathcal{K}, t \in [T] \\ y_{j,t}^c \leq C y_j^f & \forall j \in [J], t \in [T] \\ \sum\limits_{k\in[J]} y_{j,k,t}^s \leq y_{j,t}^c & \forall j \in [J], t \in [T] \\ \boldsymbol{y}^f \in \{0,1\}^J, \boldsymbol{y}^c \in \mathbb{Z}_+^{JT}, \boldsymbol{y}^s \in \mathbb{R}_+^{J^2T} & \end{array} \right\},$$

where the first constraint dictates that the vaccine stock allocated to group k population must exceed the requirement. The rest enforces inventory constraints, where C is a large constant.

4.2 Robust epidemiological optimization model

This section introduces a stochastic compartmental model designed to capture the dynamics' inherent uncertainty. Following this, we present a risk measure—ambiguity tolerance measure, which can be seamlessly integrated into a risk-aware robust optimization framework.

4.2.1 Stochastic compartmental model

Given the volatile nature of disease spread, the number of individuals in each compartment of an epidemiological model can vary widely over time. While traditional deterministic models offer some insights, they fail to account for the inherent risks and uncertainties that are crucial in medical resource planning. For example, when hospital bed occupancy approaches full capacity, waiting times can spike significantly in stochastic settings with unpredictable patient arrivals—details that deterministic models tend to overlook. Moreover, even when deterministic models suggest the same average rate of arrivals, the risk profile can differ substantially.

To recover the risks related to disease propagation, let us consider the stochastic compartmental models. A stochastic compartmental model describes the populations as random variables and forecasts the distribution of these random variables. We use $\tilde{Z}_{j,t}^n$, $n \in \mathcal{K}, j \in [J], t \in [T]$ to represent the random populations in group j and compartment n at period t, which are stochastic processes that depend on the various parameters and intervention decisions, $\boldsymbol{x}$. Based on binomial distributions (see *e.g.,* Lekone and Finkenstädt, 2006), the state dynamics of the stochastic epidemiological model can be captured as follows:

$$\tilde{Z}_{j,t+1}^n \sim \sum_{m\in\mathcal{K}/\{n\}} x_{j,t}^{m,n} + \sum_{m\in\mathcal{K}} \mathrm{Bin}\left(\tilde{Z}_{j,t}^m - \sum_{m'\in\mathcal{K}/\{m\}} x_{j,t}^{m,m'}, \hat{q}_{j,t}^{m,n} \right) \quad \forall\, n \in \mathcal{K}, j \in [J], t \in [T-1], \tag{4.5}$$

where $\mathrm{Bin}(\tilde{v}, p)$ denotes a composite binomial distribution with success probability p over $\tilde{v}$ independent Bernoulli trials. The dynamics in (4.5) describe the complex joint distributions of population sizes in the various compartments and groups over the horizon.

For convenience, we use $\hat{\mathbb{P}}$ to denote the corresponding joint distribution characterized by (4.5). The stochastic epidemiological model converges in expectation to the deterministic model with the order $O(1/N)$ as shown in Lemma 4.1. The result will also be used in the proof of Theorem 4.5.

Lemma 4.1. *Given a finite T,*

$$\lim_{N_j \to +\infty} \mathbb{E}_{\hat{\mathbb{P}}} \left| \frac{\tilde{Z}^m_{j,t} - \bar{Z}^m_{j,t}}{N_j} \right| = 0, \quad \forall t \in [T],\ m \in \mathcal{K},\ j \in [J]$$

Hence, the stochastic epidemiological model is a natural extension of the deterministic epidemiological model and incorporates risks in the model.

4.2.2 Ambiguity tolerance measure

The deterministic optimization model is oblivious to risk and uncertainty; hence, it can underestimate the efforts and costs needed to control the actual effect of an epidemic below the allotted healthcare capacities. For instance, the operational constraints $f_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y})$ are not robust to adversarial realizations of the compartment sizes that deviate from $\bar{\mathbf{Z}}$. Incorporating the stochastic epidemiological dynamics (4.5) into the optimization model makes it possible to account for the risk in disease propagation. Under the stochastic modeling context, the operational constraints should be satisfied under certain risk levels. Suppose we focus on the size of an infected population, represented by the random variable $\tilde{z} \sim \hat{\mathbb{P}}$, where $\hat{\mathbb{P}}$ is a given distribution. Suppose there is a capacity on the size of this infected population, H, we want to enforce due to healthcare considerations. Under the stochastic setting, we may have the constraint $\tilde{z} \leq H$ almost surely, requiring the number of infected cases to be less than the healthcare capacity. This constraint is fully robust and can be overly conservative. A less conservative but still quite robust and computationally tractable metric is preferred. Moreover, the distribution $\hat{\mathbb{P}}$ modeled by the stochastic epidemiological model would likely deviate from the true distribution, $\mathbb{P}^*$. The probability guarantees under $\hat{\mathbb{P}}$ do not necessarily translate to any guarantee under $\mathbb{P}^*$. A metric that can address risk and uncertainty is desirable.

To address the above issues, we propose using the riskiness index of Aumann and Serrano (2008) to evaluate the constraint violation. Let $\mathcal{L}$ be the space of all bounded discrete random variables. For a given random variable $\tilde{v} \in \mathcal{L}$ with discrete probability distribution, $\hat{\mathbb{P}}$, $\tilde{v} \sim \hat{\mathbb{P}}$ and support

$$\mathcal{V} = \left\{ v \in \mathbb{R} \mid \hat{\mathbb{P}} \left\{ \tilde{v} = v \right\} > 0 \right\},$$

we define the *ambiguity tolerance*, which is synonymous to the Aumann and Serrano (2008) riskiness index, as the functional, $\rho : \mathcal{L} \mapsto [0, \infty]$,

$$\rho\left[\tilde{v}\right] := \min\left\{\theta \mid \mu_\theta\left[\tilde{v}\right] \leq 0, \theta \geq 0\right\},$$

where $\mu_\theta : \mathcal{L} \mapsto \mathbb{R}$, $\theta \in [0, \infty]$ is the certainty equivalent under the exponential disutility as follows

$$\mu_\theta\left[\tilde{v}\right] := \begin{cases} \theta \log \mathbb{E}_{\hat{\mathbb{P}}}\left[\exp\left(\tilde{v}/\theta\right)\right] & \text{if } \theta \in (0, \infty) \\ \mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{v}\right] & \text{if } \theta = \infty \\ \max\{v \mid v \in \mathcal{V}\} & \text{if } \theta = 0. \end{cases}$$

This measure was introduced in Section 3.1.2 of the previous chapter as the entropic risk measure.

In addition to the properties of μ_θ discussed in the previous chapter, the ambiguity tolerance has the following properties with respect to uncertainty (*e.g.*,parameter misspecification).

Theorem 4.1 (Bound and Probability Guarantees)**.** *Let $\tilde{v} \in \mathcal{L}$ be a discrete random variable with support $\mathcal{V}$ over the distribution $\hat{\mathbb{P}}$ and $\theta = \rho[\tilde{v}]$. Suppose the true distribution $\mathbb{P}^*$ is absolutely continuous with respect to $\hat{\mathbb{P}}$ then*

1.

$$\mathbb{E}_{\mathbb{P}^\star}\left[\tilde{v}\right] \leq \mu_\theta\left[\tilde{v}\right] + \theta\phi_{KL}(\mathbb{P}^*\|\hat{\mathbb{P}})$$

where $\phi_{KL}(\mathbb{Q}\|\hat{\mathbb{P}})$ is the KL-divergence as follows

$$\phi_{KL}(\mathbb{Q}\|\hat{\mathbb{P}}) = \sum_{v \in \mathcal{V}} \left(\mathbb{Q}\left\{\tilde{v} = v\right\} \log\left(\frac{\mathbb{Q}\left\{\tilde{v} = v\right\}}{\hat{\mathbb{P}}\left\{\tilde{v} = v\right\}}\right)\right).$$

2. For all $a \geq 0$

$$\mathbb{P}^\star\left\{\tilde{v} > \theta\left(a + \Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right)\right)\right\} \leq \exp(-a),$$

where

$$\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right) = \max_{v \in \mathcal{V}} \log\left(\frac{\mathbb{P}^\star\left\{\tilde{v} = v\right\}}{\hat{\mathbb{P}}\left\{\tilde{v} = v\right\}}\right).$$

By Theorem 4.1, the true expected value of $\tilde{v}$ under $\mathbb{P}^*$, $\mathbb{E}_{\mathbb{P}^\star}\left[\tilde{v}\right]$, is bounded by $\mu_\theta\left[\tilde{v}\right] + \theta\phi_{KL}(\mathbb{P}^*\|\hat{\mathbb{P}})$. The violation of the upper estimate diminishes with the product of the ambiguity tolerance and how far the true distribution, $\mathbb{P}^*$, may deviate from the estimated distribution, $\hat{\mathbb{P}}$. Note that any parameter misspecification would lead to deviations of $\mathbb{P}^*$ from the estimated distribution $\hat{\mathbb{P}}$. Therefore, we can obtain robustness under parameter misspecification. Moreover, ambiguity tolerance can provide a meaningful probability guarantee when the true distribution is not observable but is proximal to the estimated distribution. Specifically, $\rho\left[\tilde{Z} - H\right]$ is associated with the probability of violations that diminishes exponentially in the magnitude of the violation as follows:

$$\mathbb{P}^{\star}\left\{\tilde{Z} > H + a\rho[\tilde{Z} - H]\right\} \leq \exp\left(-\left(a - \Delta\left(\mathbb{P}^{*}||\hat{\mathbb{P}}\right)\right)\right),$$

for all $a > \Delta\left(\mathbb{P}^{*}||\hat{\mathbb{P}}\right)$. The probability bound decreases exponentially. The strength of the bound also depends on how close the estimated distribution is with respect to the true distribution. If $\mathbb{P}^{*}$ and $\hat{\mathbb{P}}$ are identical distributions, then $\Delta\left(\mathbb{P}^{*}||\hat{\mathbb{P}}\right) = 0$ and we have for all $a > 0$,

$$\mathbb{P}^{\star}\left\{\tilde{Z} > H + a\rho[\tilde{Z} - H]\right\} \leq \exp(-a).$$

Lower ambiguity tolerance is consistent with a lower risk of capacity violation, even under an ambiguous distribution. If this ambiguity tolerance is zero, there can be no violation of the capacity.

Based on the construction of the ambiguity tolerance measure, $\rho\left(\tilde{Z} - \mu_{\bar{\rho}}\left[\tilde{Z}\right]\right) \leq \bar{\rho}$ for a given ambiguity tolerance $\bar{\rho}$. In other words, the risk of the actual compartment size, $\tilde{Z}$, exceeding the forecast, $\mu_{\bar{\rho}}\left[\tilde{Z}\right]$, is smaller than $\bar{\rho}$. We propose to take $\mu_{\bar{\rho}}[\tilde{Z}]$ as a robust epidemiological forecast of the compartment size $\tilde{Z}$. This would guarantee $\rho(\tilde{Z} - H) \leq \bar{\rho}$ for any $H \geq \mu_{\bar{\rho}}\left[\tilde{Z}\right]$. According to Theorem 4.1, the bound on the ambiguity tolerance measure would further indicate guarantees on the probability and magnitude of the capacity violation. For details on how the ambiguity tolerance controls the violation, refer to the toy example (Example 3.2).

4.2.3 Robust optimization model

We now elucidate how we can employ a robust epidemiological optimization model to help hedge against risks and uncertainty in satisfying the operational constraints. As an extension to the deterministic model of Problem (4.2), for a given acceptable cost target $\tau \geq Z_0$, we propose a *robust* model that lexicographically minimizes (see, *e.g.,* Isermann, 1982; Qi, 2017) the ambiguity tolerances, which is widely used to ensure the flexibility in handling multiple criteria and provide prioritization to recognize the importance of different criteria in decision-making. The robust epidemiological optimization model can be written as follows:

$$\begin{array}{lll}
\text{lexmin} & \boldsymbol{\theta} & \\
\text{s.t.} & \mu_{\theta_0}\left[\sum_{n\in\mathcal{K}} \mathbf{C}^n \bullet \tilde{\mathbf{Z}}^n + \boldsymbol{c}_1^\top \boldsymbol{x} + \boldsymbol{c}_2^\top \boldsymbol{y}\right] \leq \tau & \\
& \text{Dynamics constraint } (4.1) & \\
& \mu_{\theta_i}\left[f_i\left(\tilde{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}\right)\right] \leq 0 & \forall i \in [I_f] \\
& \hat{q}_{j,t}^{m,n} = \eta_{j,t}^{m,n}(\bar{\mathbf{Z}}, \boldsymbol{x}) & \forall (m,n) \in \mathcal{D}, j \in [J], t \in [T-1] \\
& \bar{\mathbf{Z}}^n \in \mathbb{R}_+^{J\times T} & \forall n \in \mathcal{K} \\
& (\boldsymbol{x}, \boldsymbol{y}) \in \bar{\mathcal{X}}. &
\end{array} \tag{4.6}$$

Minimizing the collective ambiguity tolerance levels via lexicographic minimization effectively reduces the risks of violating healthcare resource capacities and epidemiological performance targets. Similar models that minimize the riskiness index have been proved to be effective (*e.g.,* Xie, et al., 2023; Zhou, et al., 2022b). The lexicographic minimization criterion is computationally accessible as we demonstrate in the solution algorithm. On the contrary, minimizing a summation of $\boldsymbol{\theta}$ is computationally demanding.

While we evaluate the transmission rate using the deterministic dynamics as before, we transfer the deterministic operational cost objective to a risk constraint and replace the deterministic operational constraints $f_i(\bar{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}) \leq 0$ with a risk-aware counterpart $\mu_{\theta_i}\left(f_i\left(\tilde{\mathbf{Z}}, \boldsymbol{x}, \boldsymbol{y}\right)\right) \leq 0$, where $\tilde{\mathbf{Z}}$ follows dynamics (4.5). For example, a robust capacity restriction constraint can be written as

$$\mu_\theta\left[\sum_{n\in\mathcal{K}_1}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\tilde{Z}_{j,t}^n - H_t^n\right] \leq 0,$$

for some $\mathcal{K}_1 \subseteq \mathcal{K}$, $\mathcal{J} \subseteq [J]$, and $\mathcal{T} \subseteq [T]$. Note that in practice, the policymakers can also set the ambiguity tolerance $\boldsymbol{\theta}$ and solve Model (4.6) with τ as the objective to obtain a feasible solution.

Compared to deterministic epidemiological optimization, we must use an efficient method to reformulate $\mu_\theta\left[\tilde{Z}\right]$. In the next section, we show how we can recursively derive $\mu_\theta\left[\tilde{Z}\right]$ when $\tilde{Z}$ follows the stochastic dynamics as described in (4.5). This involves evaluating the associated certainty equivalents of the random population sizes at the previous time epoch scaled by a constant.

4.3 Solution approach

In this section, we focus on the reformulation of the robust epidemiological optimization model. To improve model tractability, we adopt mild approximations of the robust optimization model.

4.3.1 Reformulation of the robust optimization model

In Theorem 4.2, we show that the certainty equivalent of any compartment in the robust epidemiological model can be calculated via recursion for a fixed ambiguity tolerance.

Theorem 4.2 (Recursion of Certainty Equivalents)**.** *Ignoring the integer requirements of the number of trials of the composite Multinomial generating function, the certainty equivalent of the random populations in different compartments can be calculated recursively via the following* *system of equations: For any* $\boldsymbol{v} \in \mathbb{R}^K$, $j \in [J], t \in [T-1]$,

$$\mu_\theta \left[\boldsymbol{v}^\top \tilde{\boldsymbol{Z}}_{j,t+1}\right] = \sum_{n \in \mathcal{K}} \sum_{m \in \mathcal{K}/\{n\}} v_n x_{j,t}^{m,n} + \mu_\theta \left[\sum_{m \in \mathcal{K}} \left(\tilde{Z}_{j,t}^m - \sum_{m' \in \mathcal{K}/\{m\}} x_{j,t}^{m,m'}\right) \kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}\right)\right], \tag{4.7}$$

where the function, $\kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}\right)$ *is defined as follows*

$$\kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}\right) := \theta \log\left(\sum_{n \in \mathcal{K}} \hat{q}_{j,t}^{m,n} \exp\left(v_n/\theta\right)\right).$$

In addition, when $\theta \to +\infty$, μ_θ becomes the expectation measure, illustrating that the robust epidemiological model is closely related to the deterministic epidemiological model. Note that $\kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}\right) \geq \left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}\right)^\top \boldsymbol{v}$, we can treat $\kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}\right)$ as incorporating a buffer on top of $\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}\right)^\top \boldsymbol{v}$. Consequently, Equation (4.7) indicates that the proposed model accounts for adversarial scenarios and protects against parameter uncertainty. Based on Theorem 4.2, we can calculate the risk-aware counterparts in Model (4.6) and reformulate the robust optimization model in Theorem 4.3.

Theorem 4.3 (Model Reformulation)**.** *The robust epidemiological optimization model* (4.6) *can be reformulated as follows:*

$$\begin{aligned}
\text{lexmin} \quad & \boldsymbol{\theta} \\
\text{s.t.} \quad & \textit{Dynamics constraint} \;\; (4.1) \\
& \sum_{\substack{n \in \mathcal{K},\\ j \in [J]}} \left(\sum_{\substack{t \in [T-1],\\ m \in \mathcal{K}/\{n\}}} x_{j,t}^{m,n}\left(v_{0,j,t+1}^n - \kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}_{0,j,t+1}\right)\right) + v_{0,j,1}^n \bar{Z}_{j,1}^n\right) + \boldsymbol{c}_1^\top \boldsymbol{x} + \boldsymbol{c}_2^\top \boldsymbol{y} \leq \tau \\
& \sum_{\substack{n \in \mathcal{K},\\ j \in [J]}} \left(\sum_{\substack{t \in [T-1],\\ m \in \mathcal{K}/\{n\}}} x_{j,t}^{m,n}\left(v_{i,j,t+1}^n - \kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}, \boldsymbol{v}_{i,j,t+1}\right)\right) + v_{i,j,1}^n \bar{Z}_{j,1}^n\right) + b_{i,1}^\top \boldsymbol{x} + b_{i,2}^\top \boldsymbol{y} \leq h_i, \forall i \in [I_f] \\
& v_{0,j,T}^n = C_{j,T}^n, && \forall n \in \mathcal{K}, j \in [J] \\
& v_{i,j,T}^n = A_{i,j,T}^n, && \forall n \in \mathcal{K}, j \in [J], i \in [I_f] \\
& \kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{n,\cdot}, \boldsymbol{v}_{0,j,t+1}\right) + C_{j,t}^n \leq v_{0,j,t}^n, && \forall n \in \mathcal{K}, j \in [J], t \in [T-1] \\
& \kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{n,\cdot}, \boldsymbol{v}_{i,j,t+1}\right) + A_{i,j,t}^n \leq v_{i,j,t}^n, && \forall n \in \mathcal{K}, j \in [J], t \in [T-1], i \in [I_f] \\
& \hat{q}_{j,t}^{m,n} = \eta_{j,t}^{m,n}(\bar{\mathbf{Z}}, \boldsymbol{x}) && \forall (m,n) \in \mathcal{D}, j \in [J], t \in [T-1] \\
& \bar{\mathbf{Z}}^n \in \mathbb{R}_+^{J \times T} && \forall n \in \mathcal{K} \\
& (\boldsymbol{x}, \boldsymbol{y}) \in \bar{\mathcal{X}},
\end{aligned} \tag{4.8}$$

where $A^n_{i,j,t}$ is the coefficient in Equation (4.3).

There is an inherent complication in the aforementioned reformulation. $\kappa_\theta\left(\hat{\boldsymbol{q}}^{n,\cdot}_{j,t}, \boldsymbol{v}_{i,j,t+1}\right)$ is convex in $\boldsymbol{v}_{i,j,t+1}$ and concave in $\hat{\boldsymbol{q}}^{n,\cdot}_{j,t}$. It is challenging to find the global optimum to Problem (4.8). Hence, we resort to a second-order approximation, namely the Gaussian approximation, detailed in the next subsection.

4.3.2 Gaussian approximation

This section delves into the approximation of $\mu_\theta(\cdot)$ aiming to enhance computational efficiency. We present the asymptotic behavior of the moment-generating function related to the random populations of compartments. In the following, we show how to calculate the variance and covariance recursively for the stochastic compartment model. This result will be used to design efficient algorithms in the later optimization problem.

Theorem 4.4 (Variance and Covariance)**.** *For any $n \in \mathcal{K},\ j \in [J]$, and $t \in [T]$, the variance term $\sigma^{n,n}_{j,t+1}$ can be calculated as the follows:*

$$\sigma^{n,n}_{j,t+1} = \sum_{m\in\mathcal{K}}\left(\sigma^{m,m}_{j,t}(\hat{q}^{m,n}_{j,t})^2 + \left(\bar{Z}^m_{j,t} - \sum_{l\in\mathcal{K}\setminus\{m\}} x^{m,l}_{j,t}\right)\hat{q}^{m,n}_{j,t}(1-\hat{q}^{m,n}_{j,t})\right) + \sum_{m\neq m'\in\mathcal{K}} \sigma^{m,m'}_{j,t}\hat{q}^{m,n}_{j,t}\hat{q}^{m',n}_{j,t}$$

For any $n' \neq n \in \mathcal{K}$, the covariance term $\sigma^{n,n'}_{j,t}$ can be calculated as the follows:

$$\sigma^{n,n'}_{j,t+1} = \sum_{m\in\mathcal{K}}\left(\sigma^{m,m}_{j,t}\hat{q}^{m,n}_{j,t}\hat{q}^{m,n'}_{j,t} - \left(\bar{Z}^m_{j,t} - \sum_{l\in\mathcal{K}\setminus\{m\}} x^{m,l}_{j,t}\right)\hat{q}^{m,n}_{j,t}\hat{q}^{m,n'}_{j,t}\right) + \sum_{m\neq m'\in\mathcal{K}} \sigma^{m,m'}_{j,t}\hat{q}^{m,n}_{j,t}\hat{q}^{m',n'}_{j,t}$$

Remark 4.1. *The equations can be represented in matrix forms.*

$$\boldsymbol{\sigma}_{j,t+1} = (\hat{\boldsymbol{q}}_{j,t})^\top \boldsymbol{\sigma}_{j,t}\hat{\boldsymbol{q}}_{j,t} + \sum_{m\in\mathcal{K}}\left(\bar{Z}^m_{j,t} - \sum_{l\in\mathcal{K}\setminus\{m\}} x^{m,l}_{j,t}\right)\left(\operatorname{diag}\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}\right) - \hat{\boldsymbol{q}}^{m,\cdot}_{j,t}(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t})^\top\right), \tag{4.9}$$

where $\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}$ is a vector with respect to index n, i.e., $\{\hat{q}^{m,n}_{j,t}\}_{n\in\mathcal{K}}$ *and $\hat{\boldsymbol{q}}_{j,t}$ is the corresponding transition matrix,* i.e., $\{\hat{q}^{m,n}_{j,t}\}_{m,n\in\mathcal{K}}$.

To support the analysis of the asymptotic behavior, we introduce the following assumption regarding the asymptotic behavior at the initial state.

Assumption 4.1. *For a large population, the initial statuses $\tilde{\boldsymbol{Z}}_{j,1}$ scale proportionally with the total population size.*

In our models, the elements of $\boldsymbol{x}$ are adaptive decisions based on total populations, denoted as $x^{m,n}_{j,t} = N_j \cdot \chi^{m,n}_{j,t}$. Given fixed adaptive decisions $\boldsymbol{\chi}$, the values of $\hat{\boldsymbol{q}}_{j,t}$ remain constant as N_j

grows. Consequently, $\boldsymbol{\sigma}_{j,t}$ scales linearly with N_j. The distributional behavior is formalized in the following theorem.

Theorem 4.5 (Asymptotic Gaussian). *Under Assumption 4.1, given a finite T and adaptive decisions $\boldsymbol{\chi}$, as $N_j \to \infty$, for any $t \in [T]$,*

$$\frac{1}{\sqrt{N_j}}\left(\tilde{\boldsymbol{Z}}_{j,t} - \bar{\boldsymbol{Z}}_{j,t}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \bar{\boldsymbol{\sigma}}_{j,t}\right),$$

where $\bar{\boldsymbol{\sigma}}_{j,t} = \boldsymbol{\sigma}_{j,t}/N_j$, with rank $K-1$. Moreover, the moment-generating function

$$MGF_t(\boldsymbol{s}) = \mathbb{E}_{\hat{\mathbb{P}}}\left[\exp\left(\sum_{j\in[J]}\left(\frac{\tilde{\boldsymbol{Z}}_{j,t} - \bar{\boldsymbol{Z}}_{j,t}}{\sqrt{N_j}}\right)^{\top}\boldsymbol{s}_j\right)\right] \longrightarrow \exp\left(\sum_{j\in[J]}\frac{1}{2}\boldsymbol{s}_j^{\top}\bar{\boldsymbol{\sigma}}_{j,t}\boldsymbol{s}_j\right).$$

Theorem 4.5 shows that any subset of compartments is asymptotically Gaussian, which supports the Gaussian approximation in reformulating the risk constraints. If $\tilde{v} \sim \mathcal{N}(\bar{v}, \bar{\sigma}^2)$, then $\mu_\theta[\tilde{v}] = \bar{v} + \frac{1}{2\theta}\bar{\sigma}^2$. When $\tilde{v}$ does not follow a normal distribution, we say $\bar{v} + \frac{1}{2\theta}\bar{\sigma}^2$ is the Gaussian approximation of $\mu_\theta(\tilde{v})$, which simply evaluates the certainty equivalent as if the random variable of interests follows a normal distribution. Robust model (4.6) under the Gaussian approximation becomes

$$\begin{aligned}
\text{lexmin}\quad & \boldsymbol{\theta} \\
\text{s.t.}\quad & \sum_{t\in[T]}\sum_{j\in[J]}\left(\boldsymbol{C}_{j,t}^{\top}\bar{\boldsymbol{Z}}_{j,t} + \frac{1}{2\theta_i}\boldsymbol{C}_{j,t}^{\top}\boldsymbol{\sigma}_{j,t}\boldsymbol{C}_{j,t}\right) + \boldsymbol{c}_1^{\top}\boldsymbol{x} + \boldsymbol{c}_2^{\top}\boldsymbol{y} \le \tau \\
& \text{Dynamics constraints (4.1) and (4.9)} \\
& \sum_{t\in[T]}\sum_{j\in[J]}\left(\boldsymbol{A}_{i,j,t}^{\top}\bar{\boldsymbol{Z}}_{j,t} + \frac{1}{2\theta_i}\boldsymbol{A}_{i,j,t}^{\top}\boldsymbol{\sigma}_{j,t}\boldsymbol{A}_{i,j,t}\right) + \boldsymbol{b}_{i,1}^{\top}\boldsymbol{x} + \boldsymbol{b}_{i,2}^{\top}\boldsymbol{y} \le h_i && \forall i \in [I_f] \\
& \hat{q}_{j,t}^{m,n} = \eta_{j,t}^{m,n}(\bar{\boldsymbol{Z}}, \boldsymbol{x}) && \forall (m,n) \in \mathcal{D}, j \in [J], t \in [T-1] \\
& \bar{\boldsymbol{Z}}^n, \boldsymbol{\sigma}^{n,n} \in \mathbb{R}_+^{J\times T} && \forall n \in \mathcal{K} \\
& (\boldsymbol{x}, \boldsymbol{y}) \in \bar{\mathcal{X}}.
\end{aligned} \tag{4.10}$$

In contrast to the deterministic model (4.2), Model (4.8) from Theorem 4.3 introduces an extra set of $(I_f+1)KJT$ variables (*i.e.*, $v_{i,j,t}^n$). In Model (4.10), however, we have a set of $(K+1)KJT$ variables (*i.e.*, $\sigma_{j,t}^{n,m}$), which are uniquely based on the SEIR-type epidemiological model's structure. If $I_f \gg K$, this suggests that the Gaussian approximation model incurs a substantially lower number of auxiliary variables, making it particularly efficient for problems laden with numerous constraints, especially those tied to compartments and temporal periods.

As mentioned in Section 4.3.1, Model (4.8) includes the constraint with a non-convex logarithmic term $\kappa_\theta\left(\hat{\boldsymbol{q}}_{j,t}^{n,:}, \boldsymbol{v}_{i,j,t+1}\right)$ when $\hat{\boldsymbol{q}}_{j,t}^{n,:}$ contains the decision-dependent components. According to

Equation (4.4), dynamics constraints (4.9) within Model (4.10) can contain trilinear terms such as $\sigma_{j,t}^{m,m'} \hat{q}_{j,t}^{m,n} \hat{q}_{j,t}^{m',n}$ under a given $\boldsymbol{\theta}$ if (m,n) and $(m',n) \in \mathcal{D}$, which can be further converted to bilinear terms. Hence, Model (4.10) is more tractable than Model (4.8). It's worth noting that the deterministic epidemiological model is inherently bilinear. Any deterministic model's solution approaches can be directly applied to solve the robust Gaussian approximation model for a specified $\boldsymbol{\theta}$. Subsequently, the Bisection method, detailed in the appendix of Fu, et al. (2022), can ascertain the lexicographically optimal $\boldsymbol{\theta}$ value. Here is a polished and more scientific version of your paragraph: Subsequently, the Bisection method, as detailed in the appendix of Fu, et al. (2022), is employed to determine the lexicographically optimal value of $\boldsymbol{\theta}$. In the same appendix, Model (4.10) is further refined by omitting the covariance components, thereby reducing the number of decision variables with only a marginal loss in the objective value. Following this refinement, a detailed comparison between the robust and deterministic models is conducted, with the results summarized in Fu, et al. (2022).

4.4 Conclusion

Effectively intervening in the impact of pandemics is a keynote for policymakers. In this paper, we generalize the deterministic SEIR-type compartmental optimization model. To yield the desired optimal outcomes under the impact of risk and uncertainty, we first capture the stochastic nature of the compartments via stochastic epidemiological dynamics. We then present an ambiguity tolerance measure that can address risk and uncertainty, which facilitates a risk-aware robust epidemiological optimization model. The robust model minimizes the collective ambiguity tolerances for violating healthcare resource constraints under ambiguity, and it can be reformulated as a deterministic non-convex optimization problem. The asymptotic Gaussian property of any subset of compartments supports adopting the Gaussian approximation when evaluating the certainty equivalent to further improve the computational accessibility of the robust model. Based on these theoretical results, we develop a Python-based toolkit that facilitates modeling and solving robust epidemiological optimization problems. A real-world numerical study demonstrates the high flexibility of our toolkit and better out-of-sample performance of robust epidemiological models than the benchmarks. Moving forward, considering the transition rate's dependence on compartment states, and delving into its variability would enhance the practicality of our robust models, which we earmark for future exploration. Moreover, a significant area for future research is the development of more efficient algorithms tailored for large-scale robust epidemiological decision problems.

Appendix

B.1 Proof of results

To simplify notations in the following proofs, we define $x_{j,t}^{m+} = \sum_{m' \in \mathcal{K}/\{m\}} x_{j,t}^{m',m}$, *i.e.*, the number of individuals moving from other compartments to compartment m; and $x_{j,t}^{m-} = \sum_{m' \in \mathcal{K}/\{m\}} x_{j,t}^{m,m'}$, *i.e.*, the number of individuals moving from compartment m to other compartments. $\mathbb{E}_t[\cdot]$ is the conditional expectation on the period t, *i.e.*, $\mathbb{E}_t[\cdot] = \mathbb{E}_{\hat{\mathbb{P}}}\left[\cdot \middle| \tilde{\boldsymbol{Z}}_t\right]$. To represent the joint distribution, we use Multi$(\cdot,\cdot)$ to denote the Multinomial distribution whose marginal distribution is Binomial.

Proof of Lemma 4.1

Proof. We use mathematical induction to prove this lemma. The result holds when $t = 1$ due to the initial condition $\bar{\boldsymbol{Z}}_1 = \tilde{\boldsymbol{Z}}_1$. Then suppose the statement holds for $t \in [T]$.

Consider the next period $t+1$, we have

$$\begin{aligned}\tilde{Z}_{j,t+1}^{m} &= x_{j,t}^{m+} + \sum_{n \in \mathcal{K}} \text{Bin}\left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}, \hat{q}_{j,t}^{n,m}\right) \\ &= x_{j,t}^{m+} + \sum_{n \in \mathcal{K}} \left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m} + \sum_{n \in \mathcal{K}} \left[\text{Bin}\left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}, \hat{q}_{j,t}^{n,m}\right) - \left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right]\end{aligned}$$

which implies that

$$\begin{aligned}&\mathbb{E}_{\hat{\mathbb{P}}}\left|\frac{\tilde{Z}_{j,t+1}^{m} - \bar{Z}_{j,t+1}^{m}}{N_j}\right| \\ &= \mathbb{E}_{\hat{\mathbb{P}}}\left|\sum_{n \in \mathcal{K}} \frac{\left(\tilde{Z}_{j,t}^{n} - \bar{Z}_{j,t}^{n}\right)\hat{q}_{j,t}^{n,m}}{N_j} + \sum_{n \in \mathcal{K}} \frac{1}{N_j}\left[\text{Bin}\left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}, \hat{q}_{j,t}^{n,m}\right) - \left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right]\right| \\ &\le \sum_{n \in \mathcal{K}} \mathbb{E}_{\hat{\mathbb{P}}}\left|\frac{(\tilde{Z}_{j,t}^{n} - \bar{Z}_{j,t}^{n})\hat{q}_{j,t}^{n,m}}{N_j}\right| + \sum_{n \in \mathcal{K}} \frac{1}{N_j}\mathbb{E}_{\hat{\mathbb{P}}}\left|\text{Bin}\left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}, \hat{q}_{j,t}^{n,m}\right) - \left(\tilde{Z}_{j,t}^{n} - x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right|,\end{aligned}$$

where the inequality is based on the triangle inequality.

$$\lim_{N_j \to \infty} \mathbb{E}_{\hat{\mathbb{P}}}\left|\frac{\left(\tilde{Z}_{j,t}^{n} - \bar{Z}_{j,t}^{n}\right)\hat{q}_{j,t}^{n,m}}{N_j}\right| = 0$$

due to our inductive assumption of period t and $\hat{q}_{j,t}^{n,m}$ is bounded by 1. For the term

$$\frac{1}{N_j}\left[\text{Bin}\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-},\hat{q}_{j,t}^{n,m}\right)-\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right],$$

we evaluate its second moment

$$\begin{aligned}
&\frac{1}{N_j^2}\mathbb{E}_{\hat{\mathbb{P}}}\left[\left(\text{Bin}\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-},\hat{q}_{j,t}^{n,m}\right)-\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right)^2\right]\\
=&\frac{1}{N_j^2}\mathbb{E}_{\hat{\mathbb{P}}}\left[\mathbb{E}_t\left[\left(\text{Bin}\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-},\hat{q}_{j,t}^{n,m}\right)-\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right)^2\right]\right]\\
=&\frac{1}{N_j^2}\mathbb{E}_{\hat{\mathbb{P}}}\left[\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}(1-\hat{q}_{j,t}^{n,m})\right]\leq\frac{1}{N_j^2}\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-}\right]\rightarrow 0,\quad \text{as } N_j \text{ increases to } +\infty.
\end{aligned}$$

This shows that $\frac{1}{N_j}\left[\text{Bin}\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-},\hat{q}_{j,t}^{n,m}\right)-\left(\tilde{Z}_{j,t}^{n}-x_{j,t}^{n-}\right)\hat{q}_{j,t}^{n,m}\right]\rightarrow 0$ in L^2- Norm, also in expectation (L^1- Norm). As a result,

$$\lim_{N\rightarrow\infty}\mathbb{E}_{\hat{\mathbb{P}}}\left|\frac{\tilde{Z}_{j,t+1}^{m}-\bar{Z}_{j,t+1}^{m}}{N_j}\right|=0,\quad \forall m\in\mathcal{K},j\in[J]$$

By mathematical induction, we complete the proof. □

Proof of Theorem 4.1

Proof. The first result has been observed in D. Z. Long, et al. (2023) using the fact (see, Föllmer and Schied, 2002) that the certainty equivalent under the exponential disutility can be expressed as $\mu_\theta[\tilde{v}]=\sup_{\mathbb{Q}\in\hat{\mathcal{P}}}\{\mathbb{E}_{\mathbb{Q}}[\tilde{v}]-\theta\phi_{KL}(\mathbb{Q}||\hat{\mathbb{P}})\}$, where $\hat{\mathcal{P}}$ is the set of all probability distributions that are absolutely continuous with respect to $\hat{\mathbb{P}}$. Note that with $\theta=\mu_\theta[\tilde{v}]>0$, we have

$$\begin{aligned}
&\mu_\theta[\tilde{v}]\leq 0\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\hat{\mathbb{P}}) && \forall\mathbb{Q}\in\hat{\mathcal{P}}\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)+\theta(\phi_{KL}(\mathbb{Q}||\hat{\mathbb{P}})-\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)) && \forall\mathbb{Q}\in\hat{\mathcal{P}}\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)+\theta\sup_{\mathbb{Q}^\dagger\in\hat{\mathcal{P}}}\left\{(\phi_{KL}(\mathbb{Q}^\dagger||\hat{\mathbb{P}})-\phi_{KL}(\mathbb{Q}^\dagger||\mathbb{P}^*))\right\} && \forall\mathbb{Q}\in\hat{\mathcal{P}}\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)+\theta\sup_{\mathbb{Q}^\dagger\in\mathcal{P}^*}\left\{(\phi_{KL}(\mathbb{Q}^\dagger||\hat{\mathbb{P}})-\phi_{KL}(\mathbb{Q}^\dagger||\mathbb{P}^*))\right\} && \forall\mathbb{Q}\in\hat{\mathcal{P}}\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)+\theta\sup_{\mathbb{Q}^\dagger\in\mathcal{P}^*}\left\{\sum_{v\in\mathcal{V}:\mathbb{P}^\star\{\tilde{v}=v\}>0}\left(\mathbb{Q}^\dagger(\tilde{v}=v)\log\left(\frac{\mathbb{P}^\star\{\tilde{v}=v\}}{\hat{\mathbb{P}}\{\tilde{v}=v\}}\right)\right)\right\} && \forall\mathbb{Q}\in\hat{\mathcal{P}}\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)+\theta\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right) && \forall\mathbb{Q}\in\hat{\mathcal{P}}\\
\Longrightarrow&\mathbb{E}_{\mathbb{Q}}[\tilde{v}]\leq\theta\phi_{KL}(\mathbb{Q}||\mathbb{P}^*)+\theta\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right) && \forall\mathbb{Q}\in\mathcal{P}^*\\
\Longrightarrow&\theta\log\mathbb{E}_{\mathbb{P}^\star}[\exp(\tilde{v}/\theta)]\leq\theta\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right)\Longrightarrow\mathbb{E}_{\mathbb{P}^\star}\left[\exp\left(\left(\tilde{v}-\theta\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right)\right)/\theta\right)\right]\leq 1,
\end{aligned}$$

where $\mathcal{P}^*$ is the set of all probability distributions being absolutely continuous with respect to $\mathbb{P}^*$. Therefore, from Chernoff bound, we have for all $a > 0$

$$\begin{aligned}\mathbb{P}^\star\left\{\tilde{v} > \theta\left(a + \Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right)\right)\right\} = & \ \mathbb{P}^\star\left\{\exp\left(\left(\tilde{v} - \theta\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right)\right)/\theta\right) > \exp(a)\right\} \\ \leq & \ \mathbb{E}_{\mathbb{P}^\star}\left[\exp\left(\left(\tilde{v} - \theta\Delta\left(\mathbb{P}^*||\hat{\mathbb{P}}\right)\right)/\theta\right)\right]\exp(-a) \leq \ \exp(-a).\end{aligned}$$

□

Proof of Theorem 4.2

Proof.

$$\begin{aligned}\mu_\theta\left[\boldsymbol{v}^\top\tilde{\boldsymbol{Z}}_{j,t+1}\right] =&\theta\log\mathbb{E}_{\mathbb{P}}\left[\exp\left(\sum_{n\in\mathcal{K}}\frac{v_n}{\theta}\tilde{Z}^n_{j,t+1}\right)\right] \\ =&\theta\log\mathbb{E}_{\mathbb{P}}\left[\prod_{m\in\mathcal{K}}\mathbb{E}_{\mathbb{P}}\left[\exp\left(\sum_{n\in\mathcal{K}}\frac{v_n}{\theta}\mathrm{Bin}\left(\tilde{Z}^m_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}\right)\right)\middle|\tilde{\boldsymbol{Z}}_t\right]\right] + \sum_{n\in\mathcal{K}}v_n x^{n+}_{j,t} \\ =&\theta\log\mathbb{E}_{\mathbb{P}}\left[\prod_{m\in\mathcal{K}}\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{1}{\theta}\boldsymbol{v}^\top\mathrm{Multi}\left(\tilde{Z}^m_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,\cdot}_{j,t}\right)\right)\middle|\tilde{\boldsymbol{Z}}_t\right]\right] + \sum_{n\in\mathcal{K}}v_n x^{n+}_{j,t} \\ =&\theta\log\mathbb{E}_{\mathbb{P}}\left[\prod_{m\in\mathcal{K}}\left(\sum_{n\in\mathcal{K}}\hat{q}^{m,n}_{j,t}\exp\left(\frac{v_n}{\theta}\right)\right)^{\tilde{Z}^m_{j,t}-x^{m-}_{j,t}}\right] + \sum_{n\in\mathcal{K}}v_n x^{n+}_{j,t} \\ =&\mu_\theta\left[\sum_{m\in\mathcal{K}}\left(\tilde{Z}^m_{j,t} - \sum_{m'\in\mathcal{K}/\{m\}}x^{m,m'}_{j,t}\right)\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}, \boldsymbol{v}\right)\right] + \sum_{n\in\mathcal{K}}\sum_{m\in\mathcal{K}/\{n\}}v_n x^{m,n}_{j,t},\end{aligned}$$

where $\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}, \boldsymbol{v}\right) := \theta\log\left(\sum_{n\in\mathcal{K}}\hat{q}^{m,n}_{j,t}\exp\left(v_n/\theta\right)\right)$. When $t = 1$, we have

$$\mu_\theta\left[\boldsymbol{v}^\top\tilde{\boldsymbol{Z}}_{j,2}\right] = \sum_{m\in\mathcal{K}}\left(\tilde{Z}^m_{j,1} - \sum_{m'\in\mathcal{K}/\{m\}}x^{m,m'}_{j,1}\right)\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,1}, \boldsymbol{v}\right) + \sum_{n\in\mathcal{K}}\sum_{m\in\mathcal{K}/\{n\}}v_n x^{m,n}_{j,1}.$$

□

Proof of Theorem 4.3

Proof. According to the definition of f_i, we have

$$\mu_{\theta_i}\left[f_i\left(\tilde{\boldsymbol{Z}}, \boldsymbol{x}, \boldsymbol{y}\right)\right] = \sum_{j\in[J]}\mu_{\theta_i}\left[\sum_{t\in[T]}\sum_{n\in\mathcal{K}}A^n_{i,j,t}\tilde{Z}^n_{j,t}\right] + \boldsymbol{b}^\top_{i,1}\boldsymbol{x} + \boldsymbol{b}^\top_{i,2}\boldsymbol{y} - h_i.$$

Then we reformulate $\mu_{\theta_i}\left[\sum_{t\in[T]}\sum_{n\in\mathcal{K}} A^n_{i,j,t}\tilde{Z}^n_{j,t}\right] \le u$ recursively. Based on Theorem 4.2,

$$
\begin{aligned}
&\mu_{\theta_i}\left[\sum_{t\in[T-1]}\sum_{n\in\mathcal{K}} A^n_{i,j,t}\tilde{Z}^n_{j,t} + \sum_{n\in\mathcal{K}} A^n_{i,j,T}\tilde{Z}^n_{j,T}\right] \le u \\
\iff &\begin{cases}
v^n_{i,j,T} = A^n_{i,j,T} \\
\mu_\theta\left[\sum_{t\in[T-1]}\sum_{n\in\mathcal{K}} A^n_{i,j,t}\tilde{Z}^n_{j,t} + \sum_{m\in\mathcal{K}}\left(\tilde{Z}^m_{j,T-1} - \sum_{m'\in\mathcal{K}/\{m\}} x^{m,m'}_{j,T-1}\right)\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,T-1}, \boldsymbol{v}_{i,j,T}\right)\right] \\
\qquad\qquad + \sum_{n\in\mathcal{K}}\sum_{m\in\mathcal{K}/\{n\}} x^{m,n}_{j,T-1} v^n_{i,j,T} \le u
\end{cases} \\
\iff &\begin{cases}
v^n_{i,j,T} = A^n_{i,j,T} \\
\sum_{n\in\mathcal{K}}\sum_{m\in\mathcal{K}/\{n\}} x^{m,n}_{j,T-1}\left(v^n_{i,j,T} - \kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,T-1}, \boldsymbol{v}_{i,j,T}\right)\right) + u_{T-1} \le u \\
\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,T-1}, \boldsymbol{v}_{i,j,T}\right) + A^m_{i,j,T-1} \le v^m_{i,j,T-1} \\
\mu_\theta\left[\sum_{t\in[T-2]}\sum_{n\in\mathcal{K}} A^n_{i,j,t}\tilde{Z}^n_{j,t} + \sum_{m\in\mathcal{K}} \tilde{Z}^m_{j,T-1} v^m_{i,j,T-1}\right] \le u_{T-1}
\end{cases}
\end{aligned}
$$

Similarly, we can show that for $\tau = T-1, ..., 2$

$$
\mu_{\theta_i}\left[\sum_{t\in[\tau-1]}\sum_{n\in\mathcal{K}} A^n_{i,j,t}\tilde{Z}^n_{j,t} + \sum_{n\in\mathcal{K}} v^n_{i,j,\tau}\tilde{Z}^n_{j,\tau}\right] \le u_\tau
$$

$$
\iff \begin{cases}
\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,\tau-1}, \boldsymbol{v}_{i,j,\tau}\right) + A^m_{i,j,\tau-1} \le v^m_{i,j,\tau-1} \\
\sum_{n\in\mathcal{K}}\sum_{m\in\mathcal{K}/\{n\}} x^{m,n}_{j,\tau-1}\left(v^n_{i,j,\tau} - \kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,\tau-1}, \boldsymbol{v}_{i,j,\tau}\right)\right) + u_{\tau-1} \le u_\tau \\
\mu_\theta\left[\sum_{t\in[\tau-2]}\sum_{n\in\mathcal{K}} A^n_{i,j,t}\tilde{Z}^n_{j,t} + \sum_{m\in\mathcal{K}} \tilde{Z}^m_{j,\tau-1} v^m_{i,j,\tau-1}\right] \le u_{\tau-1}
\end{cases}
$$

When $\tau = 2$, the inequality in the last line is equivalent to

$$
\mu_\theta\left[\sum_{m\in\mathcal{K}} \tilde{Z}^m_{j,1} v^m_{i,j,1}\right] = \sum_{m\in\mathcal{K}} \bar{Z}^m_{j,1} v^m_{i,j,1} \le u_1,
$$

because the initial states $\hat{\boldsymbol{Z}}_{j,1}$ are constants. In summary, $\mu_{\theta_i}\left[f_i\left(\tilde{\boldsymbol{Z}},\boldsymbol{x},\boldsymbol{y}\right)\right]\leq 0$ is equivalent to

$$\begin{cases} v^n_{i,j,T}=A^n_{i,j,T}, & \forall n\in\mathcal{K}, j\in[J]\\ \kappa_\theta\left(\hat{\boldsymbol{q}}^{n,\cdot}_{j,t},v^n_{i,j,t+1}\right)+A^n_{i,j,t}\leq v^n_{i,j,t}, & \forall t\in[T-1],n\in\mathcal{K},j\in[J]\\ \sum\limits_{t\in[T-1]}\sum\limits_{j\in[J]}\sum\limits_{n\in\mathcal{K}}\sum\limits_{m\in\mathcal{K}/\{n\}}x^{m,n}_{j,t}\left(v^n_{i,j,t+1}-\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t},\boldsymbol{v}_{i,j,t+1}\right)\right)+\sum\limits_{j\in[J]}\sum\limits_{n\in\mathcal{K}}v^n_{i,j,1}\bar{Z}^n_{j,1}+b^\top_{i,1}\boldsymbol{x}+b^\top_{i,2}\boldsymbol{y}\leq h_i \end{cases}$$

Similarly, $\mu_{\theta_0}\left[\sum_{n\in\mathcal{K}}\boldsymbol{C}^n\bullet\tilde{\boldsymbol{Z}}^n+\boldsymbol{c}_1^\top\boldsymbol{x}+\boldsymbol{c}_2^\top\boldsymbol{y}\right]\leq\tau$ is equivalent to the follows:

$$\begin{cases} v^n_{0,j,T}=C^n_{j,T}, & \forall n\in\mathcal{K}, j\in[J]\\ \kappa_\theta\left(\hat{\boldsymbol{q}}^{n,\cdot}_{j,t},\boldsymbol{v}_{0,j,t+1}\right)+C^n_{j,t}\leq v^n_{0,j,t}, & \forall n\in\mathcal{K},j\in[J],t\in[T-1]\\ \sum\limits_{n\in\mathcal{K}}\sum\limits_{j\in[J]}\left(\sum\limits_{t\in[T-1]}\sum\limits_{m\in\mathcal{K}/\{n\}}x^{m,n}_{j,t}\left(v^n_{0,j,t+1}-\kappa_\theta\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t},\boldsymbol{v}_{0,j,t+1}\right)\right)+v^n_{0,j,1}\bar{Z}^n_{j,1}\right)+\boldsymbol{c}_1^\top\boldsymbol{x}+\boldsymbol{c}_2^\top\boldsymbol{y}\leq\tau, \end{cases}$$

□

Proof of Theorem 4.4

Proof. We first consider the variance term of the compartment n,

$$\begin{aligned}\sigma^{n,n}_{j,t+1}&=\mathrm{Var}_{\hat{\mathbb{P}}}\left[\sum_{m\in\mathcal{K}}\mathrm{Bin}\left(\tilde{Z}^m_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)\right]\\ &=\sum_{m\in\mathcal{K}}\mathrm{Var}_{\hat{\mathbb{P}}}\left(\mathrm{Bin}\left(\tilde{Z}^m_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)\right)\\ &\quad+\sum_{m\neq m'\in\mathcal{K}}\mathrm{cov}_{\hat{\mathbb{P}}}\left(\mathrm{Bin}\left(\tilde{Z}^m_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right),\mathrm{Bin}\left(\tilde{Z}^{m'}_{j,t}-x^{m'-}_{j,t},\hat{q}^{m',n}_{j,t}\right)\right).\end{aligned}$$

Applying the law of total variance, we have

$$\begin{aligned}&\mathrm{Var}_{\hat{\mathbb{P}}}\left(\mathrm{Bin}\left(\tilde{Z}^m_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)\right)\\ &=\mathrm{Var}_{\hat{\mathbb{P}}}\left[\left(\tilde{Z}^m_{j,t}-x^{m-}_{j,t}\right)\hat{q}^{m,n}_{j,t}\right]+\mathbb{E}_{\hat{\mathbb{P}}}\left[Var_{\hat{\mathbb{P}}}\left[\mathrm{Bin}\left(\tilde{Z}^m_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)\middle|\tilde{\boldsymbol{Z}}_t\right]\right]\\ &=\sigma^{m,m}_{j,t}(\hat{q}^{m,n}_{j,t})^2+\left(\bar{Z}^m_{j,t}-\sum_{l\in\mathcal{K}\backslash\{m\}}x^{m,l}_{j,t}\right)\hat{q}^{m,n}_{j,t}(1-\hat{q}^{m,n}_{j,t}).\end{aligned}$$

Applying the law of total covariance, we have

$$
\begin{aligned}
&\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m'}-x_{j,t}^{m'-},\hat{q}_{j,t}^{m',n}\right)\right)\\
&=\mathbb{E}_{\hat{\mathbb{P}}}\left[\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m'}-x_{j,t}^{m'-},\hat{q}_{j,t}^{m',n}\right)\middle|\tilde{\boldsymbol{Z}}_t\right)\right]\\
&\quad+\operatorname{cov}_{\hat{\mathbb{P}}}\left[\mathbb{E}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right)\middle|\tilde{\boldsymbol{Z}}_t\right),\mathbb{E}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m'}-x_{j,t}^{m'-},\hat{q}_{j,t}^{m',n}\right)\middle|\tilde{\boldsymbol{Z}}_t\right)\right]\\
&=\operatorname{cov}_{\hat{\mathbb{P}}}\left(\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-}\right)\hat{q}_{j,t}^{m,n},\ \left(\tilde{Z}_{j,t}^{m'}-x_{j,t}^{m'-}\right)\hat{q}_{j,t}^{m',n}\right)=\hat{q}_{j,t}^{m,n}\hat{q}_{j,t}^{m',n}\sigma_{j,t}^{m,m'},
\end{aligned}
$$

where in the second line, the conditional covariance in the first term is equal to zero because the two random variables (RVs) are conditionally independent (*i.e.*, given $\tilde{\boldsymbol{Z}}_t$).

Now we consider the covariance term between compartments n and n' ($n \neq n'$),

$$
\begin{aligned}
\sigma_{j,t+1}^{n,n'}&=\operatorname{cov}_{\hat{\mathbb{P}}}\left(\sum_{m\in\mathcal{K}}\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\sum_{m'\in\mathcal{K}}\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m'}-x_{j,t}^{m'-},\hat{q}_{j,t}^{m',n'}\right)\right)\\
&=\sum_{m\in\mathcal{K}}\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n'}\right)\right)\\
&\quad+\sum_{m\neq m'\in\mathcal{K}}\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m'}-x_{j,t}^{m'-},\hat{q}_{j,t}^{m',n'}\right)\right)
\end{aligned}
$$

Applying the law of total covariance, we have

$$
\begin{aligned}
&\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n'}\right)\right)\\
&=\mathbb{E}_{\hat{\mathbb{P}}}\left[\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right),\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n'}\right)\middle|\tilde{\boldsymbol{Z}}_t\right)\right]\\
&\quad+\operatorname{cov}_{\hat{\mathbb{P}}}\left(\mathbb{E}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n}\right)\middle|\tilde{\boldsymbol{Z}}_t\right),\mathbb{E}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-},\hat{q}_{j,t}^{m,n'}\right)\middle|\tilde{\boldsymbol{Z}}_t\right)\right)\\
&=\mathbb{E}_{\hat{\mathbb{P}}}\left[-\left(\tilde{Z}_{j,t}^{m}-x_{j,t}^{m-}\right)\hat{q}_{j,t}^{m,n}\hat{q}_{j,t}^{m,n'}\right]+\operatorname{cov}_{\hat{\mathbb{P}}}\left(\tilde{Z}_{j,t}^{m}\hat{q}_{j,t}^{m,n},\tilde{Z}_{j,t}^{m}\hat{q}_{j,t}^{m,n'}\right)\\
&=-\left(\bar{Z}_{j,t}^{m}-\sum_{l\in\mathcal{K}\setminus\{m\}}x_{j,t}^{m,l}\right)\hat{q}_{j,t}^{m,n}\hat{q}_{j,t}^{m,n'}+\sigma_{j,t}^{m,m}\hat{q}_{j,t}^{m,n}\hat{q}_{j,t}^{m,n'},
\end{aligned}
$$

where in the second line, the two RVs in the first term come from the same compartment $\tilde{Z}_{j,t}^{m}$; hence their joint conditional distribution is a multinomial distribution, and the conditional covariance is calculated according to the multinomial distribution. By the law of total covariance, we have

$$\begin{aligned}
&\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right),\operatorname{Bin}\left(\tilde{Z}^{m'}_{j,t}-x^{m'-}_{j,t},\hat{q}^{m',n'}_{j,t}\right)\right)\\
&=\mathbb{E}_{\hat{\mathbb{P}}}\left[\operatorname{cov}_{\hat{\mathbb{P}}}\left(\operatorname{Bin}\left(\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right),\operatorname{Bin}\left(\tilde{Z}^{m'}_{j,t}-x^{m'-}_{j,t},\hat{q}^{m',n'}_{j,t}\right)\middle|\tilde{\boldsymbol{Z}}_t\right)\right]\\
&\quad+\operatorname{cov}_{\hat{\mathbb{P}}}\left(\mathbb{E}_t\left(\operatorname{Bin}\left(\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)\right),\mathbb{E}_t\left(\operatorname{Bin}\left(\tilde{Z}^{m'}_{j,t}-x^{m'-}_{j,t},\hat{q}^{m',n'}_{j,t}\right)\right)\right)=\sigma^{m,m'}_{j,t}\hat{q}^{m,n}_{j,t}\hat{q}^{m',n'}_{j,t},
\end{aligned}$$

where the conditional covariance in the second line is zero since the two RVs are conditionally independent (*i.e.*, given $\tilde{\boldsymbol{Z}}_t$).

□

Proof of Theorem 4.5

Proof. To simplify notation, we define $\chi^{m+}_{j,t}=\sum_{m'\in\mathcal{K}/\{m\}}\chi^{m',m}_{j,t}$, $\chi^{m-}_{j,t}=\sum_{m'\in\mathcal{K}/\{m\}}\chi^{m,m'}_{j,t}$. Also, we use $\mathcal{N}(0,\bar{\boldsymbol{\sigma}}_{j,t})$ to refer to the Multivariate Gaussian distribution with covariance matrix $\bar{\boldsymbol{\sigma}}_{j,t}$ even if it is not full rank. Define $\zeta^n_{j,t}=\bar{Z}^n_{j,t}/N_j$.

We use mathematical induction to prove this theorem. When $t=1$, we have the initial condition $\tilde{Z}^n_{j,1}=\hat{Z}^n_{j,1}$ and $\bar{\boldsymbol{\sigma}}_{j,1}=\mathbf{0}$. Thus, the result holds at time $t=1$.

Assume $\frac{1}{\sqrt{N_j}}\left(\tilde{\boldsymbol{Z}}_{j,t}-\bar{\boldsymbol{Z}}_{j,t}\right)\xrightarrow{d}\mathcal{N}\left(\mathbf{0},\bar{\boldsymbol{\sigma}}_{j,t}\right)$ at time t. Consider the next time period $t+1$.

$$\begin{aligned}
\frac{\tilde{Z}^n_{j,t+1}-\bar{Z}^n_{j,t+1}}{\sqrt{N_j}}&=\sum_{m\in\mathcal{K}}\frac{\operatorname{Bin}\left(\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)-\left(\bar{Z}^{m}_{j,t}-x^{m-}_{j,t}\right)\hat{q}^{m,n}_{j,t}}{\sqrt{N_j}}\\
&=\sum_{m\in\mathcal{K}}\left[\frac{\operatorname{Bin}\left(\bar{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)-\left(\bar{Z}^{m}_{j,t}-x^{m-}_{j,t}\right)\hat{q}^{m,n}_{j,t}}{\sqrt{N_j}}+\frac{\hat{q}^{m,n}_{j,t}(\tilde{Z}^m_{j,t}-\bar{Z}^m_{j,t})}{\sqrt{N_j}}\right.\\
&\quad\left.+\frac{\operatorname{Bin}\left(\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)-\tilde{Z}^m_{j,t}\hat{q}^{m,n}_{j,t}-\operatorname{Bin}\left(\bar{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)+\bar{Z}^m_{j,t}\hat{q}^{m,n}_{j,t}}{\sqrt{N_j}}\right]
\end{aligned}$$

In the following, we analyze the above three terms inside the summation separately. For the first term, as N_j increases to $+\infty$,

$$\frac{\operatorname{Bin}\left(\bar{Z}^{m}_{j,t}-x^{m-}_{j,t},\hat{q}^{m,n}_{j,t}\right)-\left(\bar{Z}^{m}_{j,t}-x^{m-}_{j,t}\right)\hat{q}^{m,n}_{j,t}}{\sqrt{N_j}}\xrightarrow{d}\mathcal{N}\left(\mathbf{0},\left(\zeta^m_{j,t}-\chi^{m-}_{j,t}\right)\hat{q}^{m,n}_{j,t}(1-\hat{q}^{m,n}_{j,t})\right).$$

Hence, we also know the asymptotic joint distribution (across all compartments) is

$$\mathcal{N}\left(\mathbf{0},\left(\zeta^m_{j,t}-\chi^{m-}_{j,t}\right)\left(\operatorname{diag}\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}\right)-\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t})^\top\right)\right).$$

The second term $\frac{1}{\sqrt{N_j}}\hat{q}^{m,n}_{j,t}(\tilde{Z}^m_{j,t}-\bar{Z}^m_{j,t})$ is a linear function of $\tilde{Z}^m_{j,t}$, and it is easy to see that it converges to a Gaussian distribution. Following the idea similar to the proof of Theorem 7.3.2

by Chung (2001), we consider a sequence of independent events:

$$ {}^{l}\tilde{\boldsymbol{O}}^{m}_{j,t} = \text{Multi}\left(1, \hat{\boldsymbol{q}}^{m,\cdot}_{j,t}\right), $$

and

$$ \text{Bin}\left(\tilde{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}\right) = \sum_{l=1}^{\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t}} {}^{l}\tilde{O}^{m,n}_{j,t}; \quad \text{Bin}(\bar{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}) = \sum_{l=1}^{\bar{Z}^{m}_{j,t}-x^{m-}_{j,t}} {}^{l}\tilde{O}^{m,n}_{j,t}. \quad \text{(B.12)} $$

Considering the third term, we first evaluate its second moment.

$$ \begin{aligned} &\mathbb{E}_{\hat{\mathbb{P}}}\left[\left(\frac{\text{Bin}\left(\tilde{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}\right) - \tilde{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t} - \text{Bin}\left(\bar{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}\right) + \bar{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t}}{\sqrt{N_j}}\right)^2\right] \\ &= \mathbb{E}_{\hat{\mathbb{P}}}\left[\mathbb{E}_t\left(\frac{\left(\text{Bin}\left(\tilde{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}\right) - \tilde{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t} - \text{Bin}\left(\bar{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,n}_{j,t}\right) + \bar{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t}\right)^2}{N_j}\right)\right] \\ &= \mathbb{E}_{\hat{\mathbb{P}}}\left[\mathbb{E}_t\left(\frac{\left(\sum_{i=\bar{Z}^{m}_{j,t}-x^{m-}_{j,t}+1}^{\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t}} {}^{i}\tilde{O}^{m,n}_{j,t} - \tilde{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t} + \bar{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t}\right)^2}{N_j}\right)\mathbb{1}\left(\tilde{Z}^{m}_{j,t} \geq \bar{Z}^{m}_{j,t}\right)\right] \\ &\quad + \mathbb{E}_{\hat{\mathbb{P}}}\left[\mathbb{E}_t\left(\frac{\left(\sum_{i=\tilde{Z}^{m}_{j,t}-x^{m-}_{j,t}+1}^{\bar{Z}^{m}_{j,t}-x^{m-}_{j,t}} {}^{i}\tilde{O}^{m,n}_{j,t} + \tilde{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t} - \bar{Z}^{m}_{j,t}\hat{q}^{m,n}_{j,t}\right)^2}{N_j}\right)\mathbb{1}\left(\tilde{Z}^{m}_{j,t} < \bar{Z}^{m}_{j,t}\right)\right] \\ &= \mathbb{E}_{\hat{\mathbb{P}}}\left[\frac{\left|\bar{Z}^{m}_{j,t} - \tilde{Z}^{m}_{j,t}\right| \hat{q}^{m,n}_{j,t}(1 - \hat{q}^{m,n}_{j,t})}{N_j}\right] \leq \mathbb{E}_{\hat{\mathbb{P}}}\left[\frac{\left|\bar{Z}^{m}_{j,t} - \tilde{Z}^{m}_{j,t}\right|}{N_j}\right], \end{aligned} $$

where the second equation is due to Equation (B.12), the third equation calculates the variance of Binomial distributions, and the last inequality is because of $\hat{q}^{m,n}_{j,t}$ being bounded by 0 and 1. Based on Lemma 4.1, $(\tilde{Z}^{m}_{j,t} - \bar{Z}^{m}_{j,t})/N_j$ converges to 0 in expectation, *i.e.*, $\lim_{N\to\infty} \mathbb{E}_{\hat{\mathbb{P}}}\left[\frac{|\bar{Z}^{m}_{j,t} - \tilde{Z}^{m}_{j,t}|}{N_j}\right] = 0$. As a result, we show that the third term converges to 0 in L^2-norm, which also applies it converges to 0 in probability and has no impact on the asymptotic distribution.

Therefore, as N_j increases to $+\infty$,

$$ \begin{aligned} \frac{1}{\sqrt{N}}\left(\tilde{\boldsymbol{Z}}_{j,t+1} - \bar{\boldsymbol{Z}}_{j,t+1}\right) &\xrightarrow{P} \sum_{m\in\mathcal{K}} \frac{\text{Multi}\left(\bar{Z}^{m}_{j,t} - x^{m-}_{j,t}, \hat{q}^{m,\cdot}_{j,t}\right)}{\sqrt{N_j}} + \frac{1}{\sqrt{N_j}}\hat{\boldsymbol{q}}^{\top}_{j,t}\left(\tilde{\boldsymbol{Z}}_{j,t} - \bar{\boldsymbol{Z}}_{j,t}\right) \\ &\xrightarrow{d} \sum_{m\in\mathcal{K}} \mathcal{N}\left(\boldsymbol{0}, (\zeta^{m}_{j,t} - \chi^{m-}_{j,t})\left(\text{diag}\left(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t}\right) - \hat{\boldsymbol{q}}^{m,\cdot}_{j,t}(\hat{\boldsymbol{q}}^{m,\cdot}_{j,t})^{\top}\right)\right) + \hat{\boldsymbol{q}}^{\top}_{j,t}\mathcal{N}(0, \bar{\boldsymbol{\sigma}}_{j,t}) \\ &= \mathcal{N}\left(\boldsymbol{0}, \bar{\boldsymbol{\sigma}}_{j,t+1}\right), \end{aligned} $$

where $\bar{\boldsymbol{\sigma}}_{j,t+1} = \hat{\boldsymbol{q}}_{j,t}^{\top}\bar{\boldsymbol{\sigma}}_{j,t}\hat{\boldsymbol{q}}_{j,t} + \sum_{m\in\mathcal{K}}\left(\zeta_{j,t}^{m} - \chi_{j,t}^{m-}\right)\left(\operatorname{diag}\left(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot}\right) - \hat{\boldsymbol{q}}_{j,t}^{m,\cdot}(\hat{\boldsymbol{q}}_{j,t}^{m,\cdot})^{\top}\right)$. Note that ${}^{l}\tilde{\boldsymbol{O}}_{j,t}^{m}, m \in \mathcal{K}$ and $\tilde{\boldsymbol{Z}}_{j,t}$ are mutually independent, and $\bar{\boldsymbol{\sigma}}_{j,t+1} = \boldsymbol{\sigma}_{j,t+1}/N_j$.

In the above, we prove the convergence in distribution. However, the convergence in distribution is not sufficient to establish the convergence of MGF. In order to prove that MGF pointwisely converges to $\exp\left(\sum_{j\in[J]}\frac{1}{2}\boldsymbol{s}_j^{\top}\bar{\boldsymbol{\sigma}}_{j,t}\boldsymbol{s}_j\right)$, we must also verify the following equation (See Theorem 5 in Kozakiewicz (1947)): for any $m \in \mathcal{K},\ s \in \mathbb{R}$,

$$\lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(\left|\tilde{Z}_{j,t}^{m} - \bar{Z}_{j,t}^{m}\right| \geq \nu\sqrt{N_j}\right)\cdot\exp(\nu s)\right] = 0. \tag{B.13}$$

Using the triangle inequality and the subadditivity of a probability measure, we have

$$\begin{aligned}
&\hat{\mathbb{P}}\left(\left|\tilde{Z}_{j,t}^{m} - \bar{Z}_{j,t}^{m}\right| \geq \nu\sqrt{N_j}\right)\\
&\leq \hat{\mathbb{P}}\left(\left|\tilde{Z}_{j,t}^{m} - x_{j,t-1}^{m+} - \sum_{n\in\mathcal{K}}(\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-})\hat{q}_{j,t-1}^{n,m}\right| + \left|x_{j,t-1}^{m+} + \sum_{n\in\mathcal{K}}(\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-})\hat{q}_{j,t-1}^{n,m} - \bar{Z}_{j,t}^{m}\right| \geq \nu\sqrt{N_j}\right)\\
&\leq \hat{\mathbb{P}}\left(\left|\tilde{Z}_{j,t}^{m} - x_{j,t-1}^{m+} - \sum_{n\in\mathcal{K}}(\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-})\hat{q}_{j,t-1}^{n,m}\right| \geq \frac{\nu\sqrt{N_j}}{2}\right) + \hat{\mathbb{P}}\left(\left|\sum_{n\in\mathcal{K}}(\tilde{Z}_{j,t-1}^{n} - \bar{Z}_{j,t-1}^{m})\hat{q}_{j,t-1}^{n,m}\right| \geq \frac{\nu\sqrt{N_j}}{2}\right)
\end{aligned}$$

When $t > 1$, we prove the following upper bound for the first term of the right-hand side above.

$$\begin{aligned}
&\hat{\mathbb{P}}\left(\left|\tilde{Z}_{j,t}^{m} - x_{j,t-1}^{m+} - \sum_{n\in\mathcal{K}}(\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-})\hat{q}_{j,t-1}^{n,m}\right| \geq \frac{\nu\sqrt{N_j}}{2}\right)\\
&= \mathbb{E}_{\hat{\mathbb{P}}}\left[\hat{\mathbb{P}}\left(\left|\sum_{n\in\mathcal{K}}\sum_{l=1}^{\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-}}\left({}^{l}\tilde{O}_{j,t}^{n,m} - \hat{q}_{j,t-1}^{n,m}\right)\right| \geq \frac{\nu\sqrt{N_j}}{2}\middle|\tilde{\boldsymbol{Z}}_{t-1}\right)\right] \qquad (B.14)\\
&\leq \mathbb{E}_{\hat{\mathbb{P}}}\left[2\exp\left(-\frac{\nu^2 N_j/2}{\sum_{n\in\mathcal{K}}\sum_{l=1}^{\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-}}(1-0)^2}\right)\right] = 2\exp\left(-\frac{\nu^2}{2(1 - \sum_{n\in\mathcal{K}}\chi_{j,t-1}^{n-})}\right).
\end{aligned}$$

For the first equation in (B.14), we use the result $\operatorname{Bin}\left(\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-}, \hat{q}_{j,t}^{n,m}\right) = \sum_{l=1}^{\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-}}{}^{l}\tilde{O}_{j,t}^{n,m}$, which implies that $\tilde{Z}_{j,t}^{m} = x_{j,t-1}^{m+} + \sum_{n\in\mathcal{K}}\sum_{l=1}^{\tilde{Z}_{j,t-1}^{n} - x_{j,t-1}^{n-}}{}^{l}\tilde{O}_{j,t}^{n,m}$. For the inequality in (B.14), we adopt Hoeffding's inequality because ${}^{l}\tilde{O}_{j,t}^{n,m}$ are mutually independent across indices n and l given m. This upper bound does not depend on $N_j \in \mathbb{Z}^+$ and we have

$$\lim_{\nu\to\infty} 2\exp\left(-\frac{\nu^2}{2(1 - \sum_{n\in\mathcal{K}}\chi_{j,t-1}^{n-})}\right)\cdot\exp(\nu s) = 0.$$

Now, we are ready to use Mathematical induction to prove Equation (B.13). When $t = 1$, $\hat{\mathbb{P}}\left(\left|\tilde{Z}_{j,t}^{m} - \bar{Z}_{j,t}^{m}\right| \geq \nu\sqrt{N_j}\right) = 0$. Therefore, Equation (B.13) is valid.

Assume (B.13) holds when $t=\tau$. Then for the next time period $\tau+1$,

$$
\begin{aligned}
&\lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(|\tilde{Z}^m_{j,\tau+1}-\bar{Z}^m_{j,\tau+1}|\geq \nu\sqrt{N_j}\right)\cdot\exp\left(\nu s\right)\right]\\
&\leq \lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(\left|\tilde{Z}^m_{j,\tau+1}-x^{m+}_{j,\tau}-\sum_{n\in\mathcal{K}}(\tilde{Z}^n_{j,\tau}-x^{n-}_{j,\tau})\hat{q}^{n,m}_{j,\tau}\right|\geq\frac{\nu\sqrt{N_j}}{2}\right)\cdot\exp\left(\nu s\right)\right]\\
&\quad+\lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(\left|\sum_{n\in\mathcal{K}}(\tilde{Z}^n_{j,\tau}-\bar{Z}^n_{j,\tau})\hat{q}^{n,m}_{j,\tau}\right|\geq\frac{\nu\sqrt{N_j}}{2}\right)\cdot\exp\left(\nu s\right)\right].
\end{aligned}
$$

(B.14) implies that

$$
\lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(\left|\tilde{Z}^m_{j,\tau+1}-x^{m+}_{j,\tau}-\sum_{n\in\mathcal{K}}(\tilde{Z}^n_{j,\tau}-x^{n-}_{j,\tau})\hat{q}^{n,m}_{j,\tau}\right|\geq\frac{\nu\sqrt{N_j}}{2}\right)\cdot\exp\left(\nu s\right)\right]=0.
$$

According to the induction assumption and the subadditivity of probability measure,

$$
\begin{aligned}
&\lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(\left|\sum_{n\in\mathcal{K}}(\tilde{Z}^n_{j,\tau}-\bar{Z}^n_{j,\tau})\hat{q}^{n,m}_{j,\tau}\right|\geq\frac{\nu\sqrt{N_j}}{2}\right)\cdot\exp\left(\nu s\right)\right]\\
&\leq\sum_{n\in\mathcal{K}}\lim_{\nu\to+\infty}\sup_{N_j\in\mathbb{Z}^+}\left[\mathbb{P}\left(\left|\tilde{Z}^n_{j,\tau}-\bar{Z}^n_{j,\tau}\right|\hat{q}^{n,m}_{j,\tau}\geq\frac{\nu\sqrt{N_j}}{2K}\right)\cdot\exp\left(\nu s\right)\right]=0.
\end{aligned}
$$

Therefore, by Mathematical induction, we show that Equation (B.13) holds and the convergence of MGF is proved.

□

5 Unified robust decision models and applications in EMS problem

The Emergency Medical Service problem presents a critical logistical challenge centered on the optimal placement of ambulances and medical facilities to ensure a predefined level of coverage within an acceptable response time threshold (Peng, et al., 2020). Due to the inherently unpredictable nature of emergency demand, this problem involves significant uncertainty. Traditional optimization models often fall short in such settings, as they typically rely on probabilistic formulations that are computationally intensive and require precise knowledge of the underlying demand distribution.

In contrast, RO provides a compelling alternative for decision-making under uncertainty, particularly valued for its computational tractability. Unlike probabilistic approaches, RO circumvents the need to evaluate quantiles of random variables—tasks that are computationally demanding and often require knowledge of the underlying distribution. The risk-sensitive frameworks introduced in Chapters 3 and 4 share similar limitations in this regard. By focusing instead on worst-case scenarios within a predefined uncertainty set, RO offers more stable performance in managing constraint violations, especially when only limited historical data are available—a crucial advantage in EMS planning.

Given a cost-based attribute function $f(\boldsymbol{x}, \boldsymbol{z})$, where $\boldsymbol{x} \in \mathcal{X}$ is the *here-and-now* decision vector within the feasible set $\mathcal{X} \subseteq \mathbb{R}^D$ and $\boldsymbol{z} \in \mathcal{Z}$ is the uncertain input vector within the support set $\mathcal{Z} \subseteq \mathbb{R}^L$. In a deterministic model, the attribute is evaluated at a specific input vector $\hat{\boldsymbol{z}} \in \mathcal{Z}$, and a constraint on the cost attribute, $f(\boldsymbol{x}, \hat{\boldsymbol{z}}) \leq \tau$, may not ensure feasibility. A robust optimization model addresses uncertainty by ensuring

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma),$$

where $\mathcal{Z}(\gamma) \subseteq \mathcal{Z}$ is the uncertainty set controlled by the budget parameter $\gamma \in [0, \infty]$, such that $\mathcal{Z}(0) = \{\hat{\boldsymbol{z}}\}$ and $\mathcal{Z}(\infty) = \mathcal{Z}$. Uncertainty sets are typically defined by the discrepancy between the actual input vector $\boldsymbol{z}$ and specified input vector $\hat{\boldsymbol{z}}$, measured by the disparity metric:

$$\mathcal{Z}(\gamma) = \{\boldsymbol{z} \in \mathcal{Z} \mid \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \leq \gamma\}.$$

The disparity metric often takes $\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) = \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|$, where the matrix $\mathbf{D} \in \mathbb{S}^L$, referred to as the deviation matrix, determines the shape of the uncertainty set, and similar to $\hat{\boldsymbol{z}}$, it must be specified by the decision maker. The choice of the norm $\|\cdot\|$ significantly impacts the computational tractability of the robust optimization problem. Seminal works on robust optimization, such as Ben-Tal and Nemirovski (1998) and El Ghaoui, et al. (1998), consider the L_2-norm, leading to quadratic formulations of the robust linear constraint. In contrast, Bertsimas and Sim (2004) focus on the L_1-norm that results in more tractable linear optimization formats.

While robust optimization can be computationally efficient, it may yield conservative solutions if the parameters of the uncertainty sets are not properly determined from data. Given historical data, the specified input $\hat{\boldsymbol{z}}$ is usually the empirical mean, while $\mathbf{D}$ corresponds to the root of the empirical covariance, regardless of the choice of norm. However, there does not seem to be an approach specifically designed to estimate $\hat{\boldsymbol{z}}$ and $\mathbf{D}$ for different norms to reduce the conservativeness of the robust optimization model. While recent advancements, such as those in I. Wang, et al. (2023), address related challenges, the absence of a systematic approach to norm-specific estimation motivates the first goal of our paper: to develop a data-driven method for estimating $\hat{\boldsymbol{z}}$ and $\mathbf{D}$.

The choice of specified input $\hat{\boldsymbol{z}}$ chiefly relies on the historical data. However, it is unavoidable that the estimated $\hat{\boldsymbol{z}}$ may deviate from the latent input vector ζ like the actual mean of unknown distributions, due to reasons such as outliers and limited or incomplete data. We term this as *estimation uncertainty*. This concept, also discussed in the pandemic planning framework of Chapter 4, is increasingly important in data-driven robust models. There are several attempts to bolster robustness against it in prescriptive analysis using limited data. In the realm of data-driven robust optimization, T. Zhu, et al. (2022) propose a joint estimation and robust optimization framework that employs the loss function of predictive models to construct uncertainty sets. Within the context of distributionally robust optimization, some works (see, *e.g.*, Bertsimas and Van Parys 2022; Esteban-Pérez and Morales 2022; Bertsimas, et al. 2023) utilize empirical conditional distributions to form the ambiguity set.

While significant recent progress has been made in data-driven robust optimization, there is no unifying framework for considering both outcome uncertainty and estimation uncertainty. Going back to the work by Bertsimas and Sim (2004), robust decisions are designed with the statistical justification that they remain feasible when the input changes, for instance, ensuring

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \quad \Longrightarrow \quad \mathbb{P}\left\{f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau\right\} \leq \varepsilon,$$

where $\varepsilon \in (0, 1)$ is some small risk threshold and $\mathbb{P}$ denotes the underlying true (but unknown) probability distribution; see also Bertsimas, et al. (2021b). Outcome uncertainty is critical in capturing potential issues of infeasibility when the input changes. However, the misspecification of predictive models on specified input $\hat{\boldsymbol{z}}$ (*i.e.*, estimation uncertainty) can lead to degraded performance on the feasibility profile if only outcome uncertainty is considered. Therefore, another goal of this work is to propose a unified framework that addresses two layers of uncertainties to enhance the robustness of decision-making processes.

To summarize, in this chapter, we build a unified framework for robust decision models that connect robust optimization and robust satisficing paradigms with two layers of uncertainties via comprehensive robust constraints. A central element of our approach is the use of disparity metrics, particularly focusing on norm-based disparity metrics. We delineate two layers of uncertainty—*outcome uncertainty* and *estimation uncertainty*—to enhance our decision model. The first layer, *outcome uncertainty*, specifies deviations of the input vector from the specified vector, while the second layer, *estimation uncertainty*, addresses deviations of the specified vector from latent input vectors, such as the actual but unobservable mean.

We introduce the notion of *Minimum Volume Confidence Set* (MVCS), a novel method to determine parameters in *outcome disparity metric* such as deviation matrices and specified vectors. Our MVCS approach minimizes the volume of uncertainty sets while maintaining statistical principles as constraints. It aligns with the works by Hong, et al. (2021), aiming for smaller volume uncertainty sets to achieve larger feasible regions and less conservative decisions. It also relates to the minimum volume ellipsoid problem that uses ellipsoids to approximate historical data (see, *e.g.*, P. Sun and Freund 2004; Mittal and Hanasusanto 2022), with applications in robust regression and outlier detection (Van Aelst and P. Rousseeuw 2009; Ahipaşaoğlu 2015). Additionally, the minimum covariance determinant (MCD) estimator enhances this coverage-based outlier detection paradigm by fitting a subset of data with the smallest covariance determinant (P. J. Rousseeuw and Hubert, 2011). In the MVCS problem, the objective function is a log-determinant function, ensuring computational tractability for any conic representable norms and penalty functions. The selection of norms and functions depends on the decision makers' preferences regarding uncertainties and can also serve as hyper-parameters for tuning. Moreover, our MVCS approach can capture the skewness of historical data using the asymmetric disparity metric, and side information can also be incorporated. To construct the *estimation disparity metric* that is necessarily different from the outcome disparity metric for the sake of distinguishing two layers of uncertainty, we recommend the systematic approach recently proposed by T. Zhu, et al. 2022. Although solving the worst-case model or general robust optimization problems can be computationally challenging, numerous approximation methods have been developed to enhance computational tractability. Therefore, this paper will not focus on the algorithmic aspects but will instead emphasize the demonstration of how the proposed problems are reformulated into (conic optimization) formats compatible with state-of-the-art commercial solvers. Processes of some key reformulations have been integrated into the algebraic modeling package, RSOME (Z. Chen and Xiong, 2023).

The main contributions of our work may be summarized as follows. [1]

1. We propose a unified framework that combines robust optimization and robust satisficing paradigms with two layers of uncertainty. This proposal, with joint uncertainties, offers a strong feasibility guarantee under true distribution (*cf.* Section 5.1). Additionally, we advocate a unifying sequential robust satisficing decision framework to further reduce conservativeness (*cf.* Section 5.1.2).

2. We introduce the MVCS approach to characterize the (asymmetrical) outcome uncertainty (*cf.* Section 5.2), possibly with historical data without any distributional assumption (*cf.* Section 5.3). The corresponding problems can be transformed into convex optimization problems. We also develop (in the online version) the integer power cone reformulation for the general L_p-norm confidence ball with rational number $p > 1$. The integer power cone can be further re-expressed via elementary rotated second-order cones.

3. Our framework is capable of integrating an affine predictive model with side information to both outcome uncertainty and estimation uncertainty (*cf.* Section 5.5).

4. We conduct a numerical study to present the encouraging performance of the MVCS approach in terms of its robustness to outliers as well as its adaptability in capturing asymmetry and incorporating side information. We also apply our framework to a resource allocation problem and showcase its competitive performance compared to the sample average approach.

The rest of this chapter is structured as follows. Sections 5.1.1 and 5.1.2 introduce the unifying framework with the comprehensive robust constraints. In Section 5.2, we formulate the MVCS problem and present analytical solutions, establishing connections to existing prediction methodologies. The data-driven MVCS approach is discussed in Section 5.3. Section 5.4 illustrates the construction of estimation uncertainty. Section 5.5 extends our framework to

[1] This chapter is based on content from the pre-print article: Lou, Z., Chen, Z., Sim, M., Xie, J., and Xiong, P. (2024). Estimation and Prediction Procedures for Unified Robust Decision Models. Available at: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4890089`.

Authorship and contribution details (relevant to this chapter's content): Zhiyuan Lou (PhD candidate, this dissertation) served as the major contributor, leading the theoretical development of the models, drafting the academic content of the pre-print, and designing/conducting all numerical experiments. Peng Xiong contributed to tool development in the RSOME package. Note that the detailed technical content of this work (as included in the pre-print) is not fully presented in this chapter, consistent with the focused scope of the dissertation.

integrate the prediction with side information. Numerical results are presented in Section 5.6, and Section 5.7 concludes this chapter.

5.1 The robust decision model

We focus on a decision model expressed in the following optimization format:

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \\ & \boldsymbol{x} \in \mathcal{X}, \end{aligned}$$

where the second element $\boldsymbol{z} \in \mathcal{Z} \subset \mathbb{R}^L$ is the input vector of model parameters whose outcome could be uncertain. Here, $\mathcal{Z}$ is referred to as the *input support set*, typically a convex set that delineates the underlying physical limits and includes provisions for unforeseen rare events or even *black swans*—scenarios that have never occurred in the past.

For simplicity, our discussion is limited to a linear cost function while maintaining the attribute function within the constraint. However, our approach can be expanded to include models with multiple attributes, some of which may also be minimized within the decision's objective function. Note that the attribute function can be articulated as a linear optimization problem:

$$f(\boldsymbol{x}, \boldsymbol{z}) = \min \left\{ \boldsymbol{d}^\top \boldsymbol{y} \mid \mathbf{B}\boldsymbol{y} \geq \mathbf{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{b}(\boldsymbol{z}),\ \boldsymbol{y} \in \mathbb{R}^{D_2} \right\},$$

where the maps $\mathbf{A}(\boldsymbol{z}) : \mathbb{R}^L \to \mathbb{R}^{M_2 \times D}$, $\boldsymbol{b}(\boldsymbol{z}) : \mathbb{R}^L \to \mathbb{R}^{M_2}$ are affinely dependent on $\boldsymbol{z}$. Here, $\boldsymbol{y}$ is called the *wait-and-see recourse vector* associated with the attribute function because whenever the input vector is uncertain, its optimal solution is only determined after the input vector materializes.

5.1.1 Uncertainty in the decision model

At one end, in a deterministic decision model, uncertainty in model parameters is typically ignored, and the input vector is captured by a specified input vector $\hat{\boldsymbol{z}} \in \mathcal{Z}$ as exemplified as follows:

$$V = \left\{ \begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \hat{\boldsymbol{z}}) \leq \tau \\ & \boldsymbol{x} \in \mathcal{X}. \end{aligned} \right. \tag{5.1}$$

However, the solution to the deterministic decision model is fragile as it cannot even ensure the constraint feasibility when model parameters deviate from the specified input vector.

On the other end, the worst-case model replaces the deterministic constraint with the worst-case constraint

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}$$

and solves the following problem

$$\bar{V} = \begin{cases} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z} \\ & \boldsymbol{x} \in \mathcal{X}. \end{cases}$$

Yet, if the worst-case cost $\bar{V}$ is significantly higher than V, the worst-case solution is arguably conservative and may be deemed unacceptable.

Disparity budget and tolerance

The above discussion raises the need to balance both ends of deterministic and worst-case decision models and, more specifically, to ensure constraint feasibility while controlling the model's conservativeness. Our first attempt introduces the following comprehensive outcome robust constraint:

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z}, \tag{5.2}$$

Here, $\gamma \in [0, \infty]$, referred to as *outcome disparity budget*, governs the size of the uncertainty set $\{\boldsymbol{z} \in \mathbb{R}^L \mid \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \leq \gamma\}$ around the specified input vector, within which feasibility of the constraint $f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau$ is ensured; while $\kappa \in [0, \infty]$, referred to as *outcome disparity tolerance*, articulates the tolerance to potential violations when the actual input vector does not reside in the uncertainty set $\{\boldsymbol{z} \in \mathbb{R}^L \mid \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \leq \gamma\}$ and by tolerating such violations, it help mitigate the conservativeness of the worst-case decision model that must obey the constraint for all $\boldsymbol{z} \in \mathcal{Z}$.

Outcome and estimation uncertainties

Lying at the core of the robust constraint (5.2) is a disparity metric $\Delta(\cdot, \cdot)$ that measures the discrepancy between the actual input vector $\boldsymbol{z}$ and specified input vector $\hat{\boldsymbol{z}}$.

Definition 5.1 (Disparity Metric)**.** A function $\Delta : \mathbb{R}^L \times \mathbb{R}^L \to [0, \infty]$ is a disparity metric if it satisfies the following properties:

- **Positive Definite**: For all $\boldsymbol{z}, \boldsymbol{\zeta} \in \mathbb{R}^L$, $\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) \geq 0$. Moreover, $\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) = 0$ if and only if $\boldsymbol{z} = \boldsymbol{\zeta}$.
- **Convex**: $\Delta(\boldsymbol{z}, \boldsymbol{\zeta})$ is a convex function in $\boldsymbol{z}, \boldsymbol{\zeta} \in \mathcal{Z}$.

A disparity metric is coherent if it also satisfies the following properties:

- **Translation Invariant**: $\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) = \Delta(\boldsymbol{z} - \boldsymbol{\xi}, \boldsymbol{\zeta} - \boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \mathbb{R}^L$.
- **Positive Homogeneous**: $\Delta(k\boldsymbol{z}, k\boldsymbol{\zeta}) = k\Delta(\boldsymbol{z}, \boldsymbol{\zeta})$ for all $k \geq 0$.

A disparity metric is symmetric if $\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) = \Delta(\boldsymbol{\zeta}, \boldsymbol{z})$ for all $\boldsymbol{z}, \boldsymbol{\zeta} \in \mathbb{R}^L$.

Equipped with the disparity metric, the robust constraint (5.2) accounts for *outcome uncertainty*—the first layer of uncertainty we identify—that pertains to how the actual input vector $\boldsymbol{z}$ may deviate from the specified input vector $\hat{\boldsymbol{z}}$.

Quite notably, we identify and emphasize a second layer of uncertainty, termed *estimation uncertainty*, which pertains to the potential discrepancies between the specified input vector and the latent input vector that we denote as $\boldsymbol{\zeta} \in \hat{\mathcal{Z}} \subseteq \mathbb{R}^L$. In this context, outcome uncertainty is now associated with how the actual input vector $\boldsymbol{z}$ may deviate from this latent input vector $\boldsymbol{\zeta}$ instead of the specified input vector $\hat{\boldsymbol{z}}$. The latent vector is typically associated with the underlying random process that generates input uncertainty, and the specified input vector $\hat{\boldsymbol{z}} \in \mathcal{Z}$ is determined through an estimation process designed to identify this latent vector. For instance, if the input uncertainty originates from a distribution whose mean is $\boldsymbol{\zeta}$, then $\hat{\boldsymbol{z}}$ might represent the mean estimated from data. It is crucial to note that while historical samples of the input vector may be available to assess outcome uncertainty, empirical data for the latent input $\boldsymbol{\zeta}$ are typically not accessible.

Throughout this paper, we will use the *outcome disparity metric* $\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}})$ to measure the discrepancy between the actual input vector $\boldsymbol{z}$ and the specified input vector $\hat{\boldsymbol{z}}$, and likewise, use the *estimation disparity metric* $\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}})$ to measure the discrepancy between the latent input vector $\boldsymbol{\zeta}$ and the specified input vector $\hat{\boldsymbol{z}}$. These disparity metrics are fundamental building blocks for developing different types of robust decision models. They hedge against uncertainty while ensuring that the solutions are not overly conservative compared to worst-case scenarios. For this propose, we also define the *outcome confidence set*

$$\mathcal{B}(\hat{\boldsymbol{z}}, \gamma) = \{\boldsymbol{z} \in \mathbb{R}^L \mid \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \leq \gamma\}$$

and the *estimation confidence set*

$$\hat{\mathcal{B}}(\hat{\boldsymbol{z}}, \hat{\gamma}) = \{\boldsymbol{\zeta} \in \mathbb{R}^L \mid \hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) \leq \hat{\gamma}\}.$$

Centered at $\hat{\boldsymbol{z}}$, the radii of these confidence sets are controlled by γ and $\hat{\gamma}$, respectively.

Comprehensive robust constraint with joint uncertainties

Consolidating disparity budget and tolerance as well as outcome and estimation uncertainty, in this paper we introduce the comprehensive robust constraint with both outcome and estimation uncertainties as follows:

$$f(\boldsymbol{x},\boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) - \gamma)^{+} + \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \hat{\gamma})^{+} \quad \forall \boldsymbol{z} \in \mathcal{Z},\, \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \tag{5.3}$$

with $\gamma, \hat{\gamma}, \kappa, \hat{\kappa} \in [0,\infty]$. Interpretations of the term $\kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) - \gamma)^{+}$ related to outcome uncertainty remain similar to that of the outcome robust constraint (5.2), except now that outcome uncertainty is associated with how the actual input vector $\boldsymbol{z}$ would deviate from this latent input vector $\boldsymbol{\zeta}$. It is noteworthy that keeping $\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})$ will reduce constraint (5.3) to constraint (5.2), thereby eliminating the buffer against estimation uncertainty. Additionally, estimation uncertainty is addressed by the term $\hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \hat{\gamma})^{+}$, wherein analogously, $\hat{\gamma}$ and $\hat{\kappa}$ respectively account for *estimation disparity budget* and *estimation disparity tolerance*, and the *estimation disparity metric* $\hat{\Delta}(\boldsymbol{\zeta},\hat{\boldsymbol{z}})$ expresses the deviation from the specified input vector $\hat{\boldsymbol{z}}$ to the latent input vector $\boldsymbol{\zeta}$. The following observations illustrate the generality of the comprehensive robust constraint (5.3).

- Whenever we have zero disparity tolerance (*i.e.*, $\kappa = \hat{\kappa} = 0$) or extreme disparity budget (*i.e.*, $\gamma = \hat{\gamma} = \infty$), the comprehensive robust constraint (5.3) would recover the worst-case constraint:
$$f(\boldsymbol{x},\boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}.$$
- The case of $\hat{\kappa} = \infty$ and $\gamma = 0$ yields the tolerated robust constraint as in Ben-Tal, et al. (2017):
$$f(\boldsymbol{x},\boldsymbol{z}) \leq \tau + \kappa \min_{\boldsymbol{\zeta} \in \hat{\mathcal{Z}} \cap \hat{\mathcal{B}}(\hat{\boldsymbol{z}},\hat{\gamma})} \Delta(\boldsymbol{z},\boldsymbol{\zeta}) \quad \forall \boldsymbol{z} \in \mathcal{Z}.$$
- When $\kappa, \hat{\kappa} = \infty$ and $\hat{\gamma} = 0$ (which would enforce $\boldsymbol{\zeta} = \hat{\boldsymbol{z}}$), the comprehensive robust constraint (5.3) would be equivalent to the following classical budgeted robust constraint:
$$f(\boldsymbol{x},\boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z} \cap \mathcal{B}(\hat{\boldsymbol{z}},\gamma);$$
see, *e.g.*, Bertsimas and Sim (2004).
- When $\kappa, \hat{\kappa} = \infty$ and $\gamma = 0$ (which would enforce $\boldsymbol{z} = \boldsymbol{\zeta}$), the comprehensive robust constraint (5.3) leads to another budgeted robust constraint:
$$f(\boldsymbol{x},\boldsymbol{\zeta}) \leq \tau \qquad \forall \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \cap \hat{\mathcal{B}}(\hat{\boldsymbol{z}},\hat{\gamma}),$$
which appears in the joint estimation and robustness optimization framework of T. Zhu, et al. (2022) that mitigates estimation uncertainty in optimization problems with the integration of the parameter estimation procedure and the optimization problem.

- More recently, the tolerance robust satisficing model described by Sim, et al. (2021) represents a special case of (5.3) with $\gamma = \hat{\gamma} = 0$. Outside of the authors' focal data-driven context, this special case can be explicitly defined as the following tolerated robust constraint:

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) + \hat{\kappa}\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathcal{Z},\ \boldsymbol{\zeta} \in \hat{\mathcal{Z}}.$$

Indeed, any $\boldsymbol{x}$ feasible to the comprehensive robust constraint (5.3) guarantees its feasibility to the following robust constraints:

$$\begin{cases} f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau & \forall \boldsymbol{z} \in \mathcal{Z} \cap \mathcal{B}(\hat{\boldsymbol{z}}, \gamma) \\ f(\boldsymbol{x}, \boldsymbol{\zeta}) \leq \tau & \forall \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \cap \hat{\mathcal{B}}(\hat{\boldsymbol{z}}, \gamma) \\ f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma)^+ & \forall \boldsymbol{z} \in \mathcal{Z} \setminus \mathcal{B}(\hat{\boldsymbol{z}}, \gamma) \\ f(\boldsymbol{x}, \boldsymbol{\zeta}) \leq \tau + \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+ & \forall \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \setminus \hat{\mathcal{B}}(\hat{\boldsymbol{z}}, \hat{\gamma}). \end{cases}$$

Beyond these feasibility realms, the comprehensive robust constraint (5.3) also possesses strong statistical justifications.

Theorem 5.1. *Consider the random variable $\tilde{\boldsymbol{z}} \sim \mathbb{P}$, $\mathbb{P} \in \mathcal{P}_0(\mathcal{Z})$, which denotes the random outcomes of the actual input vector. Suppose $\boldsymbol{x}$ is feasible to the comprehensive robust constraint* (5.3)*. Then the following infeasibility likelihood profile holds:*

$$\mathbb{P}\left\{f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r + \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+\right\} \leq \mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\boldsymbol{\zeta}, \gamma + r)\right\} \quad \forall r \geq 0,\ \boldsymbol{\zeta} \in \hat{\mathcal{Z}}.$$

Theorem 5.1 implies that

$$\begin{cases} \mathbb{P}\left\{f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau\right\} \leq \mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\boldsymbol{\zeta}, \gamma)\right\} & \forall \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \cap \hat{\mathcal{B}}(\hat{\boldsymbol{z}}, \hat{\gamma}) \\ \mathbb{P}\left\{f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r\right\} \leq \mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\boldsymbol{\zeta}, \gamma + r)\right\} & \forall r \geq 0,\ \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \cap \hat{\mathcal{B}}(\hat{\boldsymbol{z}}, \hat{\gamma}). \end{cases}$$

At a greater deviation, $\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) > \hat{\gamma}$, a greater constraint violation moderated by $\hat{\kappa}$ is necessary to ensure the same infeasibility probability bound. It should be noted that establishing the probability bound depends on knowledge of the true distribution. Even when the distribution is known, relying on a generic probability bound based on the likelihood of infeasibility within the outcome confidence set may overestimate the actual probability of infeasibility.

This is a generic bound that holds true for all attribute functions. As a result, the established probabilistic guarantees tend to be conservative. Better bounds can be obtained for specific attribute functions by imposing assumptions on the distributions, which we aim to avoid. However, this issue can be addressed through robust satisficing decision models, where the modeler can directly control the level of conservativeness by specifying performance targets.

5.1.2 A unifying sequential robust satisficing decision framework

Armed with the comprehensive robust constraint (5.3) parameterized by $\gamma, \kappa, \hat{\gamma}, \hat{\kappa}$—budgets and tolerances to outcome and estimation uncertainties, our unifying decision framework follows the spirit of robust satisficing (see, *e.g.*, Ben-Tal, et al. 2017, D. Z. Long, et al. 2023) to optimize these parameters sequentially, rather than pre-specifying them as fixed. Specifically, we advocate solving sequential robust satisficing problems with *intermediate* targets as follows.

Step 0. Solve the deterministic decision model and obtain its objective value V to be used as the reference for setting the intermediate targets for the subsequent robust satisficing models.

Step 1. Solve a budgeted robust satisficing problem for the first intermediate target $V_1 \geq V$:

$$\begin{aligned}
\max \quad & \gamma \\
\text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \qquad \forall \boldsymbol{z} \in \mathcal{Z} \cap \{\boldsymbol{z} \in \mathbb{R}^L \mid \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \leq \gamma\} \\
& \boldsymbol{c}^\top \boldsymbol{x} \leq V_1 \\
& \boldsymbol{x} \in \mathcal{X}, \ \gamma \geq 0,
\end{aligned}$$

and obtain the optimal $\gamma^\star$.

Step 2. Solve a tolerated robust satisficing model that minimizes κ for the second intermediate target $V_2 \geq V_1$:

$$\begin{aligned}
\min \quad & \kappa \\
\text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma^\star)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z} \\
& \boldsymbol{c}^\top \boldsymbol{x} \leq V_2 \\
& \boldsymbol{x} \in \mathcal{X}, \ \kappa \geq 0,
\end{aligned}$$

and obtain its optimal $\kappa^\star$.

Step 3. Solve a budgeted robust satisficing model that maximizes $\hat{\gamma}$ for the third intermediate target $V_3 \geq V_2$:

$$\begin{aligned}
\max \quad & \hat{\gamma} \\
\text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa^\star(\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) - \gamma^\star)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z}, \ \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \cap \{\boldsymbol{\zeta} \in \mathbb{R}^L \mid \hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) \leq \hat{\gamma}\} \\
& \boldsymbol{c}^\top \boldsymbol{x} \leq V_3 \\
& \boldsymbol{x} \in \mathcal{X}, \ \hat{\gamma} \geq 0,
\end{aligned}$$

and obtain its optimal $\hat{\gamma}^\star$.

Step 4. Solve a robust satisficing model that minimizes $\hat{\boldsymbol{\kappa}}$ for the forth intermediate target $V_4 \geq V_3$:

$$\begin{aligned} \min \quad & \hat{\kappa} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa^\star(\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) - \gamma^\star)^+ + \hat{\kappa}(\hat{\Delta}(\hat{\boldsymbol{z}}, \boldsymbol{\zeta}) - \hat{\gamma}^\star)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z},\ \boldsymbol{\zeta} \in \hat{\mathcal{Z}} \\ & \boldsymbol{c}^\top \boldsymbol{x} \leq V_4 \\ & \boldsymbol{x} \in \mathcal{X},\ \hat{\kappa} \geq 0, \end{aligned}$$

and obtain its optimal solution $\boldsymbol{x}^\star$.

This approach of solving sequential robust satisficing models for tuning multiple parameters is attributed to Sim, et al. (2021). The approach starts by solving a deterministic optimization problem to establish the minimum feasible cost V. This initialization presumes the specified input vector $\hat{\boldsymbol{z}}$ accurately mirrors the actual (uncertain) input vector. Progressively, the approach optimizes budgets and tolerances to outcome and estimation uncertainties. The optimization occurs sequentially with increasing intermediate target $V_4 \geq V_3 \geq V_2 \geq V_1$ relative to V. This structured, incremental approach effectively addresses outcome and estimation uncertainties, informing a robust, comprehensive decision-making solution. It also simplifies parameter specification, offering decision-makers a clear and manageable progression through the model's complexities. Nevertheless, decision-makers can always articulate their budgets and tolerances to outcome and estimation uncertainties by specifying the value of any parameter among $\gamma, \hat{\gamma}, \kappa, \hat{\kappa}$. When doing so, they should skip the corresponding step among Steps 1-4.

Note that in the *budgeted robust satisficing* phase (Steps 1 and 3), disparity budgets γ and $\hat{\gamma}$ are maximized, typically by a bisection search due to the quasiconcave objective; whereas in the *tolerated robust satisficing* phase (Steps 2 and 4), the tolerance parameters κ and $\hat{\kappa}$ are minimized through a convex optimization problem. It is also worth noting that the framework is readily applied to the reward context wherein the attribute function $f(\boldsymbol{x}, \boldsymbol{z})$ is reward-based. Accordingly, the comprehensive robust constraint shall read as

$$f(\boldsymbol{x}, \boldsymbol{z}) \geq \tau - \kappa(\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) - \gamma)^+ - \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+ \quad \forall \boldsymbol{z} \in \mathcal{Z},\ \boldsymbol{\zeta} \in \hat{\mathcal{Z}}.$$

An illustration from portfolio optimization

Using portfolio optimization in the reward context as an example, we illustrate the rich expressiveness of the sequential robust satisficing decision framework. In the portfolio optimization problem, the random stock returns, denoted by $\tilde{\boldsymbol{z}} \sim \mathbb{P}$, $\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^L)$, follows an ambiguous distribution with an unknown true mean and covariance $\hat{\boldsymbol{\Sigma}} \in \mathbb{S}^L_{++}$. Suppose first that the sample mean well estimates the true mean so the specified input vector $\hat{\boldsymbol{z}}$ is aligned with the sample mean $\hat{\boldsymbol{\mu}}$. Consider the investment decision $\boldsymbol{x} \in \mathbb{R}^L$ residing in $\mathcal{X} = \{\boldsymbol{x} \in \mathbb{R}^L \mid \mathbf{1}^\top \boldsymbol{x} = 1\}$. The tolerated robust satisficing model for portfolio optimization has the following representation:

$$\kappa^\star = \begin{cases} \min & \kappa \\ \text{s.t.} & \boldsymbol{x}^\top \boldsymbol{z} \geq V - \kappa \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathbb{R}^L \\ & \boldsymbol{x} \in \mathcal{X}, \ \kappa \geq 0, \end{cases}$$

which corresponds to Step 2 of the framework with $\mathcal{Z} = \mathbb{R}^L$ that specifies $\gamma^\star = 0$ and skips Step 1. Here, V can be viewed as some risk-free return. Suppose the outcome disparity metric is $\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) = \|\hat{\boldsymbol{\Sigma}}^{-1/2}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|_2$. Using standard techniques in robust optimization to transform the constraint

$$\boldsymbol{x}^\top \boldsymbol{z} \geq V - \kappa \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathbb{R}^L,$$

the representation can be expressed as follows:

$$\kappa^\star = \begin{cases} \min & \sqrt{\boldsymbol{x}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{x}} \\ \text{s.t.} & \boldsymbol{x}^\top \hat{\boldsymbol{\mu}} \geq V \\ & \boldsymbol{x} \in \mathcal{X}, \end{cases} \tag{5.4}$$

which turns out to be the celebrated Markowitz portfolio model that minimizes the portfolio's standard deviation.

We next extend to consider estimation uncertainty where the sample mean does not coincide with the true mean and the latent vector $\boldsymbol{\zeta}$ is aligned with the true mean, *i.e.*, $\mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}] = \boldsymbol{\zeta}$. In executing the sequential robust satisficing decision framework, we can continue with the solution of the Markowitz portfolio model (5.4) for an intermediate target of V_2, and solve for the following tolerated robust satisficing model—Step 4 of the framework that specifies $\hat{\gamma}^\star = 0$ and skips Step 3—to achieve a desired return target of $V_4 \leq V_2$:

$$\begin{aligned} \min \quad & \hat{\kappa} \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{z} \geq V_4 - \kappa^\star \Delta(\boldsymbol{z}, \boldsymbol{\zeta}) - \hat{\kappa} \hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z}, \ \boldsymbol{\zeta} \in \mathbb{R}^L \\ & \boldsymbol{x} \in \mathcal{X}, \ \hat{\kappa} \geq 0. \end{aligned}$$

Suppose the estimation disparity metric is $\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) = \|\boldsymbol{\zeta} - \hat{\boldsymbol{z}}\|_2$. We can express the above model as

$$\begin{aligned} \min \quad & \|\boldsymbol{x}\|_2 \\ \text{s.t.} \quad & \boldsymbol{x}^\top \hat{\boldsymbol{z}} \geq V_4 \\ & \sqrt{\boldsymbol{x}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{x}} \leq \kappa^\star \\ & \boldsymbol{x} \in \mathcal{X}. \end{aligned} \tag{5.5}$$

A well-known empirical study by DeMiguel, et al. (2009) has demonstrated that an equal-weighted portfolio $\boldsymbol{x} = \mathbf{1}/L$ can potentially outperform the Markowitz portfolio. Interestingly, if the equal-weighted portfolio has an expected return exceeding V_4 and a standard deviation lower than $\kappa^\star$, then it would emerge as an optimal solution to Problem (5.5). That being said, by adjusting V_2 and V_4 within the range of $V_4 \leq V_2$, the nature of the optimal portfolio can shift, ranging from the solution proposed by the Markowitz portfolio to an equal-weighted portfolio.

5.2 Characterizing outcome uncertainty via minimum volume confidence set

In this section, we explore methods to specify parameters that will characterize the outcome uncertainty within the comprehensive robust constraint. Our primary focus is on the norm-based disparity metric:

$$\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) = \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|,$$

where $\|\cdot\|$ represents an *absolute norm.*

Definition 5.2 (Absolute Norm)**.** An absolute norm satisfies $\|\boldsymbol{u}\| = \|\boldsymbol{v}\|$ for all $\boldsymbol{u} \in \mathbb{R}^L$, with $v_\ell = |u_\ell|$ for each $\ell \in [L]$. The properties of the absolute norm are detailed in Bertsimas and Sim (2006). In particular, for any $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{R}^L$ such that $v_\ell \geq |u_\ell|$ for all $\ell \in [L]$, $\|\boldsymbol{v}\| \geq \|\boldsymbol{u}\|$. Examples of absolute norms include the L_p-norm and the D-norm introduced in Bertsimas, et al. (2004).

For the norm-based disparity metric, we define the matrix $\mathbf{D} \in \mathbb{S}^L_{++}$ as the *deviation matrix*, which is both symmetric and positive definite. This matrix plays two crucial roles: it adjusts to the unique deviations of each input entry and ensures the outcome disparity metric is dimensionless. These features are vital for fair and consistent evaluations across different input entries, especially when they differ in units, scales, or inherent variability. Consequently, we redefine the outcome confidence set to include the dependency of the deviation matrix as follows:

$$\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma) = \{\boldsymbol{z} \in \mathbb{R}^L \mid \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\| \leq \gamma\}.$$

The radius of the confidence set is controlled by the disparity budget γ, and its volume is affected by both γ and the deviation matrix $\mathbf{D}$ (as we shall present subsequently).

Infeasibility likelihood profile and deviation penalty function

The *infeasibility likelihood profile* φ establishes the upper bound for the infeasibility probability of the outcome confidence set as follows:

$$\mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)\right\} \leq \varphi(\gamma) \quad \forall \gamma \geq 0.$$

To see a concrete example of the infeasibility likelihood profile, consider a special case of $\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma) = \{\boldsymbol{z} \in \mathbb{R}^L \mid \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|_2 \leq \gamma\}$. Suppose the true mean $\mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}]$ and variance $\boldsymbol{\Sigma} = \mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})^\top\right]$ are known so that we can set $\hat{\boldsymbol{z}} = \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}]$ and $\mathbf{D}^2 = \boldsymbol{\Sigma}$. It holds that

$$\mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)\right\} \leq \mathbb{P}\left\{\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_2^2 \geq \gamma^2\right\} \leq \frac{\mathbb{E}_{\mathbb{P}}\left[\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_2^2\right]}{\gamma^2} = \varphi(\gamma) \quad \forall \gamma \geq 0,$$

where the second inequality follows from Markov inequality and the last equality is satisfied for an infeasibility likelihood profile defined as $\varphi(r) = L/r^2$ $\forall r \geq 0$, established based on the insight that

$$\mathbb{E}_{\mathbb{P}}\left[1/\varphi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_2)\right] = \mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})^\top \boldsymbol{\Sigma}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})/L\right] = 1.$$

The infeasibility likelihood profile serves as a key tool in the statistical justification of the comprehensive robust counterpart. For example, from Theorem 5.1, if a solution $\boldsymbol{x} \in \mathcal{X}$ is feasible to the following comprehensive robust constraint in (5.3) with $\hat{\kappa} = \infty$ and $\hat{\gamma} = 0$:

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa\left(\|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\| - \gamma\right)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z},$$

then it also conforms to the tolerated chance constraint:

$$\mathbb{P}\left\{f(\boldsymbol{x}, \boldsymbol{z}) > \tau + \kappa r\right\} \leq \mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma + r)\right\} \leq \varphi(\gamma + r) \quad \forall r \geq 0.$$

To demonstrate how the specified input vector $\hat{\boldsymbol{z}}$ and the deviation matrix $\mathbf{D}$ contribute to establishing the infeasibility bound of the outcome confidence set (thereby making it adaptable to various infeasibility likelihood profiles), we next introduce the *deviation penalty function.*

Definition 5.3 (Deviation Penalty Function)**.** Associated with an infeasibility likelihood profile φ, the *deviation penalty function* $\psi : \mathbb{R}_+ \to [0, \infty]$ is a non-negative, non-decreasing and convex function such that $\psi(0) < 1$ and satisfies

$$\psi(r) \geq \frac{1}{\varphi(r)} \quad \forall r \geq 0.$$

Specifically, for some $p \in [1, \infty]$, a *power deviation penalty function* is defined as follows:

$$\psi_p(r) = \frac{r^p}{L},$$

which is normalized by the uncertainty's dimension L.

Theorem 5.2. *Consider a deviation penalty function ψ associated with the infeasibility likelihood profile φ. Suppose the random variable $\tilde{\boldsymbol{z}} \sim \mathbb{P}$, $\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^L)$ satisfies*

$$\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \le 1,$$

for some $\hat{\boldsymbol{z}} \in \mathbb{R}^L$ and $\mathbf{D} \in \mathbb{S}^L_{++}$. Then, the infeasibility probability bound of the outcome confidence set satisfies:

$$\mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)\right\} \le \varphi(\gamma) \quad \forall \gamma \ge 0.$$

Minimum volume confidence set

Given an infeasibility likelihood profile φ and its related deviation penalty function ψ, Theorem 5.2 relates the infeasibility probability bound of an outcome confidence set $\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)$ to its parameters $\hat{\boldsymbol{z}}$ and $\mathbf{D}$. Among those values of $(\hat{\boldsymbol{z}}, \mathbf{D})$ that satisfy $\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \le 1$, we seek the ones that correspond to the smallest volume of the outcome confidence set $\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)$, thereby avoid overly conservative scenarios and solutions.

Note that changing variables from $\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})$ to $\boldsymbol{v}$, the volume of $\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)$ can be precisely calculated as:

$$\begin{aligned} \mathrm{Vol}(\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)) &= \int_{\{\boldsymbol{z}:\|\mathbf{D}^{-1}(\boldsymbol{z}-\hat{\boldsymbol{z}})\|\le\gamma\}} \mathrm{d}\boldsymbol{z} = \int_{\{\boldsymbol{v}:\|\boldsymbol{v}\|\le\gamma\}} |\det(\mathbf{D})| \mathrm{d}\boldsymbol{v} = |\det(\mathbf{D})| \cdot \mathrm{Vol}(\mathcal{B}(\mathbf{0}, \mathbf{I}, \gamma)) \\ &\propto |\det(\mathbf{D})|. \end{aligned}$$

That being said, the determinant of the deviation matrix $\mathbf{D}$ affects the volume of the outcome confidence set, which in turn, impacts the conservativeness of the comprehensive robust counterpart. Therefore, among all parameters that meet the infeasibility likelihood profile as ensured in Theorem 5.2, we should select those that minimize the determinant of the deviation matrix. This can be established as the following *minimum volume confidence set* (MVCS) problem:

$$\begin{aligned} \inf \quad & \det(\mathbf{D}) \\ \text{s.t.} \quad & \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \le 1 \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L, \ \mathbf{D} \in \mathbb{S}^L_{++}. \end{aligned} \tag{MVCS}$$

We assume that an optimal solution to the MVCS problem exists, contingent upon the covariance matrix of the random variable being positive definite. Additionally, if the specified input is determined by another method, it can be excluded as a decision variable from the MVCS problem.

Analytical solutions for MVCS

The MVCS problem is introduced as a conceptual optimization problem aimed at determining the least conservative parameters for the comprehensive robust constraint that would achieve its desired statistical performance guarantees. Solving MVCS involves precise knowledge of the true distribution $\mathbb{P}$ of $\tilde{\boldsymbol{z}}$ and evaluation of the associated expectation constraint that could be computationally challenging. Nonetheless, we can identify scenarios where closed-form solutions to the MVCS problem are obtainable. In this section, we explore such scenarios.

Symmetric Distribution When the true distribution $\mathbb{P}$ is symmetric, the optimal solution of the specified input vector $\hat{\boldsymbol{z}}$ corresponds to the random variable's mean.

Theorem 5.3. *Suppose the optimal solution to the* MVCS *problem exists and the distribution* $\mathbb{P}$ *of* $\tilde{\boldsymbol{z}}$ *is symmetric about its mean* $\boldsymbol{\mu} = \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}]$. *Then* $\hat{\boldsymbol{z}}^\star = \boldsymbol{\mu}$ *constitutes part of an optimal solution to the* MVCS *problem.*

L_p-Norm and Power Deviation Penalty Function Focusing on the L_2-norm and the corresponding power deviation penalty function ψ_2, we derive an analytical solution to the following representation of the MVCS problem that applies to any distribution $\mathbb{P}$.

$$\begin{array}{ll} \inf & \det(\mathbf{D}) \\ \text{s.t.} & \mathbb{E}_{\mathbb{P}}\left[\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_2^2\right] \leq L \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \mathbf{D} \in \mathbb{S}^L_{++}. \end{array} \tag{5.6}$$

Theorem 5.4. *Let* $\boldsymbol{\mu} = \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}]$ *and* $\boldsymbol{\Sigma} = \mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}} - \boldsymbol{\mu})(\tilde{\boldsymbol{z}} - \boldsymbol{\mu})^\top\right]$ *and suppose* $\boldsymbol{\Sigma} \in \mathbb{S}^L_{++}$. *Then an optimal solution to Problem* (5.6) *is* $(\boldsymbol{z}^\star, \mathbf{D}^\star) = (\boldsymbol{\mu}, \boldsymbol{\Sigma}^{1/2})$.

Gaussian Distribution In the case of a Gaussian distribution, the deviation matrix is a scalar multiple of the square root of the covariance matrix.

Theorem 5.5. *Suppose the distribution* $\mathbb{P}$ *of* $\tilde{\boldsymbol{z}}$ *is a Gaussian with covariance* $\boldsymbol{\Sigma} \in \mathbb{S}^L_{++}$. *Then the optimal solution to the* MVCS *problem with* L_p*-norm and power deviation penalty function* ψ_p *is* $\boldsymbol{z}^\star = \boldsymbol{\mu}$ *and*

$$\mathbf{D}^\star = \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}\right)^{1/p}\boldsymbol{\Sigma}^{1/2},$$

where $\Gamma(r) = \int_0^\infty t^{r-1}e^{-t}\mathrm{d}t$ *is the gamma function. For* $p = 2$, $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$.[2]

[2] As expected, when $p = 2$, Theorem 5.5 gives the same optimal solution as in Theorem 5.4.

Asymmetrical outcome

Accounting for asymmetry can be particularly useful in applications like portfolio management, where skewness in returns can significantly influence investment decisions. We propose the following asymmetric outcome disparity metric that is based on an absolute norm:

$$\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) = \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^{-} + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^{+}\|,$$

where $\|\cdot\|$ is an absolute norm, and $\underline{\boldsymbol{\sigma}}$ and $\bar{\boldsymbol{\sigma}}$ are asymmetric normalization vectors with positive entries. When $\underline{\boldsymbol{\sigma}} = \bar{\boldsymbol{\sigma}} = \mathbf{1}$, the property of an absolute norm and the identity that $v^- + v^+ = |v|$ yield the following:

$$\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}\boldsymbol{u})^{-} + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}\boldsymbol{u})^{+}\| = \|(\mathbf{D}^{-1}\boldsymbol{u})^{-} + (\mathbf{D}^{-1}\boldsymbol{u})^{+}\| = \|\mathbf{D}^{-1}\boldsymbol{u}\|.$$

Therefore, we have $\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) = \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|$, which aligns with the definition of symmetric disparity metric.

Accordingly, the asymmetric outcome confidence set is defined as

$$\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma) = \left\{\boldsymbol{z} \in \mathbb{R}^L \mid \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^{-} + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^{+}\| \le \gamma\right\}. \tag{5.7}$$

Note that the first asymmetric uncertainty set for robust optimization was proposed in X. Chen, et al. (2007), which is a special case of the asymmetric outcome confidence set with $\mathbf{D} = \mathbf{I}$. It is based on the assumption that the input parameters are independently distributed, sub-Gaussian, and characterized by their directional deviation vectors $\underline{\boldsymbol{\sigma}}$, $\bar{\boldsymbol{\sigma}}$, and mean $\hat{\boldsymbol{z}}$. In our approach, the parameters $\underline{\boldsymbol{\sigma}}$, $\bar{\boldsymbol{\sigma}}$, $\hat{\boldsymbol{z}}$ and $\mathbf{D}$ are all estimated from data.

Theorem 5.6. *The asymmetric outcome confidence set can be expressed as follows:*

$$\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma) = \left\{ \boldsymbol{z} \in \mathbb{R}^L \;\middle|\; \begin{array}{l} \exists \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}} \in \mathbb{R}^L_+ \; : \\ \bar{\boldsymbol{w}} - \underline{\boldsymbol{w}} = \mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}) \\ \|\,\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\underline{\boldsymbol{w}} + \mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\bar{\boldsymbol{w}}\| \le \gamma \end{array} \right\}. \tag{5.8}$$

We next evaluate the volume of the asymmetric outcome confidence set.

Theorem 5.7. *The volume of the asymmetric outcome confidence set* $\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma)$ *can be calculated as follows:*

$$\mathrm{Vol}(\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma)) = \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \cdot \mathrm{Vol}(\mathcal{B}(\mathbf{0}, \mathbf{I}, \mathbf{1}, \mathbf{1}, \gamma)) \propto \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}).$$

Hence, the MVCS problem can be extended to the asymmetric outcome confidence set as follows:

$$\begin{array}{ll} \inf & \prod_{\ell\in[L]} \dfrac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \\ \text{s.t.} & \mathbb{E}_{\mathbb{P}}\left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^+\|)\right] \leq 1 \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}} \in \mathbb{R}^L_{++},\ \mathbf{D} \in \mathbb{S}^L_{++}. \end{array}$$

5.3 Estimating minimum volume confidence set from data

The MVCS problem provides a method for estimating the specified input vector $\hat{\boldsymbol{z}}$ and the deviation matrix $\mathbf{D}$, by utilizing the empirical distribution $\hat{\mathbb{P}}$ constructed from a set of historical data $\mathcal{D} = \{(\omega, \boldsymbol{z}_\omega) \mid \omega \in [\Omega]\}$ as follows:

$$\hat{\mathbb{P}}\left\{\tilde{\boldsymbol{z}} = \boldsymbol{z}_\omega\right\} = \frac{1}{\Omega} \quad \forall \omega \in [\Omega].$$

Here, $\boldsymbol{z}_\omega$ represents the observed outcome of the input vector in scenario ω. Accordingly, the empirical mean and covariance of $\tilde{\boldsymbol{z}}$ can be concisely presented as:

$$\hat{\boldsymbol{\mu}} = \mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{z}}\right] = \frac{1}{\Omega}\sum_{\omega\in[\Omega]} \boldsymbol{z}_\omega \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{\mu}})(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{\mu}})^\top\right] = \frac{1}{\Omega}\sum_{\omega\in[\Omega]} (\boldsymbol{z}_\omega - \hat{\boldsymbol{\mu}})(\boldsymbol{z}_\omega - \hat{\boldsymbol{\mu}})^\top.$$

This method is known as the *minimum volume confidence set estimation* approach, which involves solving the following convex optimization problem to estimate $\hat{\boldsymbol{z}}$ and $\mathbf{D}$:

$$\begin{array}{ll} \inf & \det(\mathbf{D}) \\ \text{s.t.} & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \leq 1 \\ & \mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r}) \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \mathbf{D} \in \mathbb{S}^L_{++}. \end{array} \tag{$\widehat{\text{MVCS}}$}$$

Here, the conic inequality $\mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r})$ represents $\mathbf{D} - \mathrm{Diag}(\boldsymbol{r}) \in \mathbb{S}^L_+$ and the pre-specified regularization vector $\boldsymbol{r} \in \mathbb{R}^L_{++} \cup \{\mathbf{0}\}$ is applied to enforce $\mathbf{D} \in \mathbb{S}^L_{++}$, ensuring the volume of the outcome confidence set remains positive. Such a treatment is deployed to address practical issues arising with the $\widehat{\text{MVCS}}$ approach that deviates from the MVCS methodology. For example, even if the true covariance, $\boldsymbol{\Sigma}$, evaluated on the true distribution $\mathbb{P}$, is positive definite, the empirical covariance $\hat{\boldsymbol{\Sigma}}$, derived from the same distribution, may not be. This subtlety can render the optimal solution to the $\widehat{\text{MVCS}}$ problem unsolvable.

We next demonstrate how the $\widehat{\text{MVCS}}$ problem can be transformed into a convex optimization problem.

Theorem 5.8. *The optimal solution* $(\hat{\boldsymbol{z}}^\star, \mathbf{D}^\star)$ *of the* $\widehat{\mathrm{MVCS}}$ *problem can be obtained from solving the following convex optimization problem:*

$$\begin{array}{ll} \max & (\det(\mathbf{Q}))^{1/L} \\ \text{s.t.} & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|)\right] \leq 1 \\ & \mathbf{Q} \preceq \mathrm{Diag}^{-1}(\boldsymbol{r}) \\ & \boldsymbol{q} \in \mathbb{R}^L, \ \mathbf{Q} \in \mathbb{S}^L_+. \end{array} \tag{5.9}$$

Specifically, suppose the optimal solution of Problem (5.9) *is* $(\boldsymbol{q}^\star, \mathbf{Q}^\star)$*. Then* $\mathbf{D}^\star = \mathbf{Q}^{\star-1}$ *and* $\hat{\boldsymbol{z}}^\star = \mathbf{D}^\star \boldsymbol{q}^\star$*. Problem* (5.9) *is strictly feasible and regularization ensures that the solution is bounded and hence achievable. In the absence of regularization (i.e.,* $\boldsymbol{r} = \mathbf{0}$*), the constraint* $\mathbf{Q} \preceq \mathrm{Diag}^{-1}(\boldsymbol{r})$ *is eliminated.*

Apart from being a convex optimization problem, another notable advantage of Problem (5.9) is that evaluating the expectation over the empirical distribution in its constraint is computationally feasible, as shown below:

$$\mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|)\right] = \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \psi(\|\mathbf{Q}\boldsymbol{z}_\omega - \boldsymbol{q}\|).$$

If the penalty function ψ and the norm $\|\cdot\|$ are tractable and conic representable, the expectation constraint can also be expressed similarly. Specifically, when using the L_p-norm with the corresponding power deviation penalty $\psi_p(r)$, this expectation constraint can be modeled using power cones for any $p \geq 1$ and becomes linear when $p = 1$. Yet, state-of-the-art commercial solvers do not support power cones well, except MOSEK. In the appendix of the online version, we show that for rational values of $p \geq 1$, the expectation constraint can be represented using integer power cones—a new notion we have developed—and subsequently expressed as second-order cone constraints. We also explain how using such integer power cones can reformulate the determinant optimization problem as a semidefinite optimization problem. These methods have been integrated into the Python-based algebraic modeling toolbox RSOME developed by Z. Chen and Xiong (2023).

The deviation penalty function ψ, as a convex function, penalizes expected deviations from $\hat{\boldsymbol{z}}$ by using the disparity metric. Thus, from an estimation perspective, when the deviation penalty function is linear, the estimated parameters typically show minimal sensitivity to data fluctuations. Conversely, with more convex penalties, such as power functions, the estimation becomes more sensitive to variations in data. When using the deviation penalty function ψ_∞—the limit of ψ_p attained as p tends to infinity, we have the following equivalence:

$$\mathbb{E}_{\hat{\mathbb{P}}}\left[\psi_\infty(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \leq 1 \quad \Longleftrightarrow \quad \|\mathbf{D}^{-1}(\boldsymbol{z}_\omega - \hat{\boldsymbol{z}})\| \leq 1 \quad \forall \omega \in [\Omega].$$

Consequently, Problem (5.9) transforms into a minimum volume confidence set problem that covers all empirical samples. For an ellipsoidal confidence set where the outcome disparity metric is based on the L_2-norm, this aligns with the well-known minimum ellipsoid covering problem—a technique widely used in data analysis, computational geometry, and optimal experimental design in statistics, as noted by Todd (2016), and is valued for their computational efficiency, as discussed by P. Sun and Freund (2004).

Separate data sets

Another practical consideration is that data may not always be collected in a coordinated manner that would allow us to form a joint empirical distribution. More commonly, we encounter separate data sets, each associated with a subset of entries of the input vector. To capture this, specifically, consider the input vector as a concatenation of I disjoint components, *i.e.*, $\boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_I)$, where $\boldsymbol{z}_i \in \mathbb{R}^{L_i}$ for each $i \in [I]$ such that $\sum_{i\in[I]} L_i = L$. Each random component $\tilde{\boldsymbol{z}}_i$ is associated with a different dataset $\mathcal{D}_i$, and we denote its empirical distribution by $\hat{\mathbb{P}}_i \in \mathcal{P}_0(\mathbb{R}^{L_i})$ with $\tilde{\boldsymbol{z}}_i \sim \hat{\mathbb{P}}_i$. In this situation, the empirical distribution $\hat{\mathbb{P}}$ for the complete random input vector $\tilde{\boldsymbol{z}} = (\tilde{\boldsymbol{z}}_1, \ldots, \tilde{\boldsymbol{z}}_I)$ is not uniquely defined, prohibiting the evaluation of $\mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right]$.

To resolve this ambiguous expectation, we can restrict the deviation matrix to a block diagonal one such that $\mathbf{D} = \mathrm{Diag}(\mathbf{D}_1, \ldots, \mathbf{D}_I)$ for some $\mathbf{D}_i \in \mathbb{S}^{L_i}_{++}$, $i \in [I]$. Using the L_p-norm with the corresponding power deviation penalty $\psi_p(r)$ and following a similar vein as outlined in Problem (5.6), we have

$$\mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] = \mathbb{E}_{\hat{\mathbb{P}}}\left[\frac{1}{L}\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_p^p\right] = \frac{1}{L}\sum_{i\in[I]} \mathbb{E}_{\hat{\mathbb{P}}_i}\left[\|\mathbf{D}_i^{-1}(\tilde{\boldsymbol{z}}_i - \hat{\boldsymbol{z}}_i)\|_p^p\right].$$

Consequently, for separate data sets, we consider the following $\widehat{\mathrm{MVCS}}$ problem:

$$\begin{array}{lll} \inf & \det(\mathrm{Diag}(\mathbf{D}_1, \ldots, \mathbf{D}_I)) & \\ \text{s.t.} & \displaystyle\sum_{i\in[I]} \mathbb{E}_{\hat{\mathbb{P}}_i}\left[\|\mathbf{D}_i^{-1}(\tilde{\boldsymbol{z}}_i - \hat{\boldsymbol{z}}_i)\|_p^p\right] \leq L & \\ & \mathbf{D}_i \succeq \mathrm{Diag}(\boldsymbol{r}_i) & \forall i \in [I] \\ & \hat{\boldsymbol{z}}_i \in \mathbb{R}^{L_i},\ \mathbf{D}_i \in \mathbb{S}^{L_i}_{++} & \forall i \in [I]. \end{array}$$

Here, $\det(\mathrm{Diag}(\mathbf{D}_1, \ldots, \mathbf{D}_I)) = \prod_{i\in[I]} \det(\mathbf{D}_i)$ and $\boldsymbol{r}_i \in \mathbb{R}^{L_i}_{++} \cup \{\mathbf{0}\}$ $\forall i \in [I]$. Following the exposition of Theorem 5.8, the optimal solution $(\hat{\boldsymbol{z}}_i^\star, \mathbf{D}_i^\star)$ for each $i \in [I]$ can be obtained from solving a convex optimization problem

$$
\begin{array}{lll}
\max & \sqrt[L]{\prod_{i\in[I]} \det(\mathbf{Q}_i)} & \\
\text{s.t.} & \sum\limits_{i\in[I]} \mathbb{E}_{\hat{\mathbb{P}}_i}\left[\|\mathbf{Q}_i\tilde{\boldsymbol{z}}_i - \boldsymbol{q}_i\|_p^p\right] \le L & \\
& \mathbf{Q}_i \preceq \mathrm{Diag}^{-1}(\boldsymbol{r}_i) & \forall i \in [I] \\
& \boldsymbol{q}_i \in \mathbb{R}^{L_i},\ \mathbf{Q}_i \in \mathbb{S}_+^{L_i} & \forall i \in [I].
\end{array}
\tag{5.10}
$$

Specifically, for each $i \in [I]$, $\mathbf{D}_i^\star = \mathbf{Q}_i^{\star -1}$ and $\hat{\boldsymbol{z}}_i^\star = \mathbf{D}_i^\star \boldsymbol{q}_i$, where $(\boldsymbol{q}_i^\star, \mathbf{Q}_i^\star)$ is the optimal solution of Problem (5.10).

For the case of $I = L$, which implies $L_i = 1$ for all $i \in L$, we can derive the closed-form solution to Problem (5.10), which can be simplified as follows:

$$
\begin{array}{ll}
\inf & \prod\limits_{i\in[L]} d_i \\
\text{s.t.} & \sum\limits_{i\in[L]} \mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i|^p / d_i^p\right] \le L \\
& \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \boldsymbol{d} \in \mathbb{R}_+^L.
\end{array}
\tag{5.11}
$$

where $\boldsymbol{d}$ corresponds to the diagonal of the deviation matrix, *i.e.*, $\mathbf{D} = \mathrm{Diag}(\boldsymbol{d})$.

Theorem 5.9. *Suppose the empirical variance of $\tilde{z}_i$ is positive for each $i \in [L]$. Then, the optimal solution to Problem* (5.11) *is, for each $i \in [L]$,*

$$
d_i^\star = \left[\mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i^\star|^p\right]\right]^{1/p},
$$

where $\hat{z}_i^\star$ satisfies the following condition:

$$
\begin{cases}
\mathbb{E}_{\hat{\mathbb{P}}_i}\left[((\tilde{z}_i - \hat{z}_i^\star)^+)^{p-1}\right] = \mathbb{E}_{\hat{\mathbb{P}}_i}\left[((\tilde{z}_i - \hat{z}_i^\star)^-)^{p-1}\right] & \text{if } p > 1 \\
\mathbb{E}_{\hat{\mathbb{P}}_i}\left[\mathrm{sign}(\tilde{z}_i - \hat{z}_i^\star)\right] = 0 & \text{if } p = 1.
\end{cases}
$$

Theorem 5.9 offers valuable insights into how the specified input vector $\hat{\boldsymbol{z}}$ and the deviation vector $\boldsymbol{d}$ are estimated under the diagonal restriction of the deviation matrix. For $p = 2$, they correspond to the empirical means and standard deviations, respectively. For $p = 1$, $\hat{\boldsymbol{z}}$ aligns with the median of $\tilde{\boldsymbol{z}}$ while $\boldsymbol{d}$ is derived from the empirical absolute deviation relative to the median. Notably, estimating parameters with lower values of p tends to be less affected by data fluctuations.

$\widehat{\text{MVCS}}$ for asymmetric confidence set

We can extend the $\widehat{\text{MVCS}}$ problem to an asymmetric confidence set as follows:

$$
\begin{aligned}
\inf \quad & \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}} \left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^+\|) \right] \leq 1 \\
& \mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r}) \\
& \hat{\boldsymbol{z}} \in \mathbb{R}^L, \ \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}} \in \mathbb{R}^L_{++}, \ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned}
\tag{5.12}
$$

Unfortunately, Problem (5.12) cannot be transformed into a convex optimization problem, complicating the pursuit of a global optimum. Nonetheless, a viable approach is to employ an alternating optimization strategy. Begin by setting the asymmetric normalization vectors as $\underline{\boldsymbol{\sigma}} = \bar{\boldsymbol{\sigma}} = \mathbf{1}$, and then solve for the solution variables $\hat{\boldsymbol{z}}$ and $\mathbf{D}$. This step can be formulated as a convex optimization problem by applying the change of variables technique described in Theorem 5.8. After determining the optimal specified input vector $\hat{\boldsymbol{z}}$ and the optimal outcome deviation matrix $\mathbf{D}$, the next step involves refining the asymmetric normalization vectors by solving the following optimization problem:

$$
\begin{aligned}
\inf \quad & \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}} \left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^+\|) \right] \leq 1 \\
& \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}} \in \mathbb{R}^L_{++}.
\end{aligned}
\tag{5.13}
$$

These two steps are iteratively repeated until a predefined termination criterion is met. With each iteration, the volume of the asymmetric confidence set decreases. Given that this volume cannot reach zero, due to constraints imposed by regularization, the process will eventually converge to a local optimum.

Problem (5.13) can also be formulated as a convex optimization problem, as demonstrated in the subsequent result.

Theorem 5.10. *The optimal solution* $(\underline{\boldsymbol{\sigma}}^\star, \bar{\boldsymbol{\sigma}}^\star)$ *of Problem* (5.13) *can be obtained from solving the following convex optimization problem:*

$$
\begin{aligned}
\max \quad & \sqrt[L]{\textstyle\prod_{\ell \in [L]} t_\ell} \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}} \left[\psi(\|(\mathrm{Diag}(\underline{\boldsymbol{v}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}(\bar{\boldsymbol{v}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^+\|) \right] \leq 1 \\
& t_\ell^2 \leq 2\underline{v}_\ell \underline{w}_\ell \qquad \forall \ell \in [L] \\
& t_\ell^2 \leq 2\bar{v}_\ell \bar{w}_\ell \qquad \forall \ell \in [L] \\
& t_\ell \geq \underline{w}_\ell + \bar{w}_\ell \qquad \forall \ell \in [L] \\
& \boldsymbol{t}, \underline{\boldsymbol{v}}, \bar{\boldsymbol{v}}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}} \in \mathbb{R}^L_{+}.
\end{aligned}
\tag{5.14}
$$

Specifically, suppose the optimal solution of Problem (5.14) *is* $(\underline{\boldsymbol{v}}^\star, \bar{\boldsymbol{v}}^\star)$*. Then, for any* $\ell \in [L]$*,* $\underline{\sigma}_\ell^\star = 1/\underline{v}_\ell^\star$ *and* $\bar{\sigma}_\ell^\star = 1/\bar{v}_\ell^\star$*. Problem* (5.14) *is strictly feasible, ensuring that the optimal solution* $\underline{\boldsymbol{v}}^\star, \bar{\boldsymbol{v}}^\star > \mathbf{0}$*. Since* $\mathbf{D} \in \mathbb{S}_{++}^L$*, the optimal solution is also bounded and attainable.*

The objective function of Problem (5.14) is characterized by a hypograph that utilizes an integer power cone, while the constraints are formulated as rotated second-order conic constraints.

5.4 Characterizing estimation uncertainty from data

To characterize estimation uncertainty, it is tempting to use the same disparity metric as used for characterizing outcome uncertainty. However, the following result cautions that it is necessary to deploy different disparity metrics to distinguish these two layers of uncertainty.

Theorem 5.11. *Suppose* Δ *and* $\hat{\Delta}$ *are identical coherent disparity metrics and* $\mathcal{Z} \subseteq \hat{\mathcal{Z}}$*. Then, the comprehensive robust constraint with joint uncertainties as defined in* (5.3) *can be simplified to the following comprehensive robust constraint with only outcome uncertainty:*

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \min\{\kappa, \hat{\kappa}\}(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma - \hat{\gamma})^+ \quad \forall \boldsymbol{z} \in \mathcal{Z}. \tag{5.15}$$

Theorem 5.11 stresses the need to use a different coherent disparity metric for characterizing estimation uncertainty. Similar to the outcome disparity metric, we can express the estimation disparity metric based on norms, such as $\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) = \|\hat{\mathbf{D}}^{-1}(\boldsymbol{\zeta} - \hat{\boldsymbol{z}})\|$ for some absolute norm $\|\cdot\|$ and deviation matrix $\hat{\mathbf{D}} \in \mathbb{S}_{++}^L$ that are different from the ones used in the outcome disparity metric; see also the previous portfolio example wherein we deploy different norms as well as different matrices $\mathbf{D}$ and $\hat{\mathbf{D}}$ to recover the Markowitz portfolio and the equal-weighted portfolio.

Alternatively, a more systematic approach for characterizing estimation uncertainty has been proposed by T. Zhu, et al. (2022) in which the estimation disparity metric, $\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}})$ is constructed from the underlying optimization problem in which the specified input vector is estimated. Typically, the specified input vector is determined by solving the following convex optimization problem:

$$\hat{\boldsymbol{z}} = \arg\min_{\boldsymbol{\zeta} \in \hat{\mathcal{Z}}} \mathscr{L}(\boldsymbol{\zeta}; \mathcal{D}),$$

where the function $\mathscr{L}(\boldsymbol{\zeta}; \mathcal{D})$, termed the *loss function*, is a convex function that reaches its *unique* minimum at the specified input vector $\hat{\boldsymbol{z}}$. In this case, it is unnecessary to perform the $\widehat{\text{MVCS}}$ approach to identify another specified input vector for outcome uncertainty. Instead, we can use this specified input vector as input to $\widehat{\text{MVCS}}$ to determine the outcome deviation matrix $\mathbf{D}$ that defines the outcome disparity metric Δ. A translation-invariant estimation disparity metric can then be derived from the estimation process in the following manner:

$$\hat{\Delta}(\boldsymbol{u}, \mathbf{0}) = \eta(\mathscr{L}(\hat{\boldsymbol{z}} + \boldsymbol{u}; \mathcal{D}) - \mathscr{L}(\hat{\boldsymbol{z}}; \mathcal{D})),$$

where η is any monotonically increasing transformation function that satisfies $\eta(0) = 0$, such as a linear function. Such an estimation disparity metric emphasizes the reliance on the loss function's behavior around its minimum to quantify uncertainty. The choice of η is not unique, provided it ensures that the estimation disparity metric maintains the convexity property. Note that the positive definite property holds because $\hat{\boldsymbol{z}}$ is the unique minimum of the loss function.

Square loss function

Given the estimation deviation matrix $\hat{\mathbf{D}} \in \mathbb{S}^L_{++}$, which can be different from the outcome deviation matrix $\mathbf{D}$, we consider the square loss function as follows:

$$\mathscr{L}(\boldsymbol{\zeta}; \mathcal{D}) = \mathbb{E}_{\hat{\mathbb{P}}}\left[\|\hat{\mathbf{D}}^{-1}(\boldsymbol{\zeta} - \tilde{\boldsymbol{z}})\|_2^2\right],$$

whose simplification proceeds as

$$\begin{aligned}
\mathbb{E}_{\hat{\mathbb{P}}}\left[\|\hat{\mathbf{D}}^{-1}(\boldsymbol{\zeta} - \tilde{\boldsymbol{z}})\|_2^2\right] &= \mathbb{E}_{\hat{\mathbb{P}}}\left[(\boldsymbol{\zeta} - \tilde{\boldsymbol{z}})^\top \hat{\mathbf{D}}^{-2}(\boldsymbol{\zeta} - \tilde{\boldsymbol{z}})\right] \\
&= \mathbb{E}_{\hat{\mathbb{P}}}\left[\boldsymbol{\zeta}^\top \hat{\mathbf{D}}^{-2}\boldsymbol{\zeta} - 2\boldsymbol{\zeta}^\top \hat{\mathbf{D}}^{-2}\tilde{\boldsymbol{z}} + \tilde{\boldsymbol{z}}^\top \mathbf{D}^{-2}\tilde{\boldsymbol{z}}\right] \\
&= \boldsymbol{\zeta}^\top \hat{\mathbf{D}}^{-2}\boldsymbol{\zeta} - 2\boldsymbol{\zeta}^\top \hat{\mathbf{D}}^{-2}\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{z}}\right] + \mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{z}}^\top \hat{\mathbf{D}}^{-2}\tilde{\boldsymbol{z}}\right].
\end{aligned}$$

Minimizing this quadratic function leads to $\hat{\boldsymbol{z}} = \mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{z}}\right]$ and results in

$$\mathscr{L}(\hat{\boldsymbol{z}}; \mathcal{D}) = -\hat{\boldsymbol{z}}^\top \hat{\mathbf{D}}^{-2}\hat{\boldsymbol{z}} + \mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{z}}^\top \hat{\mathbf{D}}^{-2}\tilde{\boldsymbol{z}}\right].$$

Consequently, the translation-invariant estimation disparity metric associated with the square loss function simplifies to

$$\begin{aligned}
\hat{\Delta}(\boldsymbol{u}, \mathbf{0}) &= \eta(\mathbb{E}_{\hat{\mathbb{P}}}\left[(\hat{\boldsymbol{z}} + \boldsymbol{u} - \tilde{\boldsymbol{z}})^\top \hat{\mathbf{D}}^{-2}(\hat{\boldsymbol{z}} + \boldsymbol{u} - \tilde{\boldsymbol{z}})\right] - \mathscr{L}(\hat{\boldsymbol{z}}; \mathcal{D})) \\
&= \eta((\hat{\boldsymbol{z}} + \boldsymbol{u})^\top \hat{\mathbf{D}}^{-2}(\hat{\boldsymbol{z}} + \boldsymbol{u}) - 2(\hat{\boldsymbol{z}} + \boldsymbol{u})^\top \hat{\mathbf{D}}^{-2}\hat{\boldsymbol{z}} + \hat{\boldsymbol{z}}^\top \hat{\mathbf{D}}^{-2}\hat{\boldsymbol{z}}) \\
&= \eta(\|\hat{\mathbf{D}}^{-1}\boldsymbol{u}\|_2^2).
\end{aligned}$$

Note that the selection of the transformation function η can be quite flexible: it may be any convex and increasing function that originates from the zero point. Specifically, selecting a power transformation function $\eta(r) = r^p$ for $p \geq 1/2$ ensures the resulting estimation disparity metric remains convex. When setting $p = 1/2$, this approach reduces to the estimation disparity metric derived from the L_2-norm.

Observe these methods do not directly facilitate the derivation of the estimation deviation matrix $\hat{\mathbf{D}}$ from the dataset. This arises because the evaluation of the loss function does not incorporate the statistical information necessary to inform the choice of the estimation deviation matrix. T. Zhu, et al. (2022) detail a maximum likelihood estimation approach based on a multivariate Gaussian distribution. In our context, this method would yield an estimation deviation matrix $\hat{\mathbf{D}}$ that corresponds to the square root of the empirical covariance. However, the method relies on the Gaussian assumption, which may not always hold in practical scenarios such as portfolio optimization. In the portfolio optimization example, the estimation deviation matrix $\hat{\mathbf{D}} = \mathbf{I}$ would explain the robustness of the equal-weighted portfolio observed in practice (see the portfolio example discussed earlier).

5.5 Prediction with side information

We now consider the situations where we have access to a collection of historical data containing side information that can be used to predict the outcome of the input vector. The data is represented as $(\boldsymbol{z}_\omega, \boldsymbol{s}_\omega) \in \mathbb{R}^L \times \mathbb{R}^S$ for $\omega \in [\Omega]$, where $\boldsymbol{s}_\omega$ corresponds to the observed vector of *side information* that possesses some predictive power for the outcome of $\boldsymbol{z}_\omega$. Define the data set with set information as $\mathcal{D} = \{(\omega, \boldsymbol{z}_\omega, \boldsymbol{s}_\omega)\}_{\omega\in[\Omega]}$. Accordingly, we also define the empirical distribution $\hat{\mathbb{P}} \in \mathcal{P}_0(\mathbb{R}^L \times \mathbb{R}^S)$ of $(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{s}}) \sim \hat{\mathbb{P}}$ via $\hat{\mathbb{P}}\{(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{s}}) = (\boldsymbol{z}_\omega, \boldsymbol{s}_\omega)\} = 1/\Omega,\ \forall \omega \in [\Omega]$.

We focus on an affine prediction map $\boldsymbol{\chi} : \mathbb{R}^S \to \mathbb{R}^L$ that maps the side information to the predicted outcome. Specifically, for any given side information vector $\boldsymbol{s} \in \mathbb{R}^S$, $\boldsymbol{\chi}$ yields the predicted outcome under an affine prediction map as follows:

$$\boldsymbol{\chi}(\boldsymbol{s}) = \hat{\boldsymbol{z}}_0 + \hat{\mathbf{Z}}\boldsymbol{s},$$

where $\hat{\boldsymbol{z}}_0 \in \mathbb{R}^L$ and $\hat{\mathbf{Z}} \in \mathbb{R}^{L\times S}$ are parameters estimated from data. Once the prediction map is established, it can be used to determine the specified input vector: at the decision-making moment when the side information vector materializes to $\hat{\boldsymbol{s}}$, the specified input vector is defined as

$$\hat{\boldsymbol{z}} = \boldsymbol{\chi}(\hat{\boldsymbol{s}}) = \hat{\boldsymbol{z}}_0 + \hat{\mathbf{Z}}\hat{\boldsymbol{s}}.$$

With the side information vector $\hat{\boldsymbol{s}}$, we can incorporate both outcome uncertainty and estimation uncertainty in the following comprehensive robust constraint:

$$f(\boldsymbol{x}, \boldsymbol{z}) \le \tau + \kappa(\Delta(\boldsymbol{z}, \boldsymbol{\zeta}_0 + \mathbf{Z}\hat{\boldsymbol{s}}) - \gamma)^+ + \hat{\kappa}(\hat{\Delta}((\boldsymbol{\zeta}_0, \mathbf{Z}), (\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}})) - \hat{\gamma})^+ \quad \forall \boldsymbol{z} \in \mathcal{Z},\ (\boldsymbol{\zeta}_0, \mathbf{Z}) \in \hat{\mathcal{Z}}.$$

Here, the estimation disparity metric measures the discrepancy between estimated parameters $(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}})$ and latent parameters $(\boldsymbol{\zeta}_0, \mathbf{Z})$. Without considering estimation uncertainty (*i.e.*, $\hat{\kappa} = \infty$ and $\hat{\gamma} = 0$), the comprehensive robust constraint would be simplified as follows:

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \boldsymbol{\chi}(\hat{\boldsymbol{s}})) - \gamma)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z},$$

which reflects that the specified input vector $\hat{\boldsymbol{z}} = \boldsymbol{\chi}(\hat{\boldsymbol{s}})$ is updated accordingly with respect to the latest realization $\hat{\boldsymbol{s}}$ of the side information.

Minimum volume confidence set estimation

Similarly, we consider the minimum volume confidence set estimation problem in the setting with side information. We can formulate the optimization problem analogous to the $\widehat{\text{MVCS}}$ problem, while replacing the specified input with $\boldsymbol{\chi}(\tilde{\boldsymbol{s}})$ as follows:

$$\begin{aligned}
\inf \quad & \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}_0 - \hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\|)\right] \leq 1 \\
& \mathbf{D} \succeq \operatorname{Diag}(\boldsymbol{r}) \\
& \hat{\boldsymbol{z}}_0 \in \mathbb{R}^L, \ \hat{\mathbf{Z}} \in \mathbb{R}^{L\times S}, \ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned} \tag{5.16}$$

Next, we demonstrate how the minimum volume confidence set estimation problem with side information can be transformed into a convex optimization problem.

Theorem 5.12. *The optimal solution* $(\hat{\boldsymbol{z}}_0^\star, \hat{\mathbf{Z}}^\star, \mathbf{D}^\star)$ *of Problem* (5.16) *can be obtained from solving the following convex optimization problem:*

$$\begin{aligned}
\sup \quad & (\det(\mathbf{Q}))^{1/L} \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q} - \mathbf{P}\tilde{\boldsymbol{s}}\|)\right] \leq 1 \\
& \mathbf{Q} \preceq \operatorname{Diag}^{-1}(\boldsymbol{r}) \\
& \boldsymbol{q} \in \mathbb{R}^L, \ \mathbf{P} \in \mathbb{R}^{L\times S}, \ \mathbf{Q} \in \mathbb{S}^L_{++}.
\end{aligned} \tag{5.17}$$

Specifically, suppose the optimal solution of Problem (5.17) *is* $(\boldsymbol{q}^\star, \mathbf{P}^\star, \mathbf{Q}^\star)$*. Then* $\mathbf{D}^\star = \mathbf{Q}^{\star -1}$, $\hat{\boldsymbol{z}}_0^\star = \mathbf{D}^\star \boldsymbol{q}^\star$*, and* $\hat{\mathbf{Z}}^\star = \mathbf{D}^\star \mathbf{P}^\star$*. In the absence of regularization (i.e.,* $\boldsymbol{r} = \mathbf{0}$*), the constraint* $\mathbf{Q} \preceq \operatorname{Diag}^{-1}(\boldsymbol{r})$ *is eliminated.*

Then, we explore scenarios in which an analytical optimal solution can be developed for the case of L_2-norm and the corresponding power deviation penalty $\psi_2(r)$.

Theorem 5.13. *The optimal solution to Problem* (5.16) *for the specific case with an L_2-norm, a power deviation penalty $\psi_2(\cdot)$, and $\boldsymbol{r} = \mathbf{0}$ can be characterized as follows:*

$$\hat{\mathbf{Z}}^\star = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1},\ \hat{\boldsymbol{z}}_0^\star = \hat{\boldsymbol{\mu}}_{\boldsymbol{z}} - \hat{\mathbf{Z}}^\star\hat{\boldsymbol{\mu}}_{\boldsymbol{s}},\ \textit{and}\ \ \mathbf{D}^\star = (\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})^{\frac{1}{2}},$$

where $\hat{\boldsymbol{\mu}}_{\boldsymbol{s}} = \mathbb{E}_{\hat{\mathbb{P}}}[\tilde{\boldsymbol{s}}]$, $\hat{\boldsymbol{\mu}}_{\boldsymbol{z}} = \mathbb{E}_{\hat{\mathbb{P}}}[\tilde{\boldsymbol{z}}]$, $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}} = \mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}})(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}})^\top\right]$, $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}}^\top = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}} = \mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{z}})(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}})^\top\right]$, *and* $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}}$ *is assumed to be positive definite.*

Estimation from regression

To quantify estimation uncertainty incorporating side information, apart from the minimum volume confidence set estimation approach, we can also employ the methodology described in T. Zhu, et al. (2022) to define the estimation disparity metric using a linear regression procedure. In this approach, the parameters $(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}}) \in \hat{\mathcal{Z}}$ are determined by solving the optimization problem:

$$(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}}) \in \arg\min_{(\boldsymbol{\zeta}_0, \mathbf{Z}) \in \hat{\mathcal{Z}}} \mathscr{L}(\boldsymbol{\zeta}_0, \mathbf{Z}; \mathcal{D}).$$

Ordinary Least Square Regression Given any estimation deviation matrix $\hat{\mathbf{D}} \in \mathbb{S}_{++}^L$, the ordinary least square (OLS) regression minimizes the square loss function as follows:

$$\min_{\boldsymbol{\zeta}_0 \in \mathbb{R}^L,\, \mathbf{Z} \in \mathbb{R}^{L \times S}} \mathscr{L}(\boldsymbol{\zeta}_0, \mathbf{Z}; \mathcal{D}) = \min_{\boldsymbol{\zeta}_0 \in \mathbb{R}^L,\, \mathbf{Z} \in \mathbb{R}^{L \times S}} \mathbb{E}_{\hat{\mathbb{P}}}\,[\|\hat{\mathbf{D}}^{-1}(\boldsymbol{\zeta}_0 + \mathbf{Z}\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{z}})\|_2^2].$$

Theorem 5.14. *Suppose the empirical covariance matrix of side information is positive definite, (i.e., $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}} \in \mathbb{S}_{++}^S$). Then the OLS estimators are*

$$\hat{\boldsymbol{z}}_0 = \hat{\boldsymbol{\mu}}_{\boldsymbol{z}} - \hat{\mathbf{Z}}\hat{\boldsymbol{\mu}}_{\boldsymbol{s}} \quad \textit{and} \quad \hat{\mathbf{Z}} = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}.$$

The estimation disparity metric associated with the square loss function and a transformation function η is translation invariant and simplifies to:

$$\hat{\Delta}((\boldsymbol{u}, \mathbf{U}), (\mathbf{0}, \mathbf{0})) = \eta(\mathscr{L}(\hat{\boldsymbol{z}}_0 + \boldsymbol{u}, \hat{\mathbf{Z}} + \mathbf{U}; \mathcal{D}) - \mathscr{L}(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}}; \mathcal{D})) = \eta(\|\hat{\mathbf{D}}^{-1}(\boldsymbol{u} + \mathbf{U}\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})\|_2^2 + \|\hat{\mathbf{D}}^{-1}\mathbf{U}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{1/2}\|_{\mathrm{F}}^2),$$

where $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}(\mathbf{A}\mathbf{A}^\top)}$ represents the Frobenius norm of the matrix $\mathbf{A}$. Selecting a power transformation function $\eta(r) = r^p$ for some $p \geq 1/2$ ensures the resulting estimation disparity metric remains convex.

Interestingly, the OLS estimators coincide with the $\widehat{\mathrm{MVCS}}$ solutions obtained in Theorem 5.13 so that they lead to the same prediction mappings. This finding also lays the groundwork for constructing the estimation disparity metric.

5.6 Numerical experiments

We first evaluate the performance of the $\widehat{\text{MVCS}}$ approach, with a particular emphasis on its robustness to outliers and its adaptability in capturing skewness (*i.e.*, asymmetry) and incorporating side information. Then, we demonstrate the efficacy of robust decision models that utilize empirically derived outcome disparity metrics in allocating emergency service resources.

5.6.1 Performance of $\widehat{\text{MVCS}}$

In this subsection, we begin by evaluating the effectiveness of $\hat{\boldsymbol{z}}$ and confidence sets estimated by the $\widehat{\text{MVCS}}$ approach, focusing on its robustness to outliers. This is then followed by experiments on skew-normal distributions with/w.o. side information.

Symmetric Confidence Set

We start with the symmetric confidence set and investigate the impact of outliers on the outcome confidence set. To visualize this, we consider a simple case where the input vector consists of two i.i.d. normal random variables following $\mathbb{N}(0, 10)$. An outlier is generated in the same way with a drift 100. Figure 5.1 examines the symmetric confidence sets estimated via different norms and penalty functions. For the cubic penalty function ($p_2 = 3$), the boundary of the confidence set shifts rightwards as the value of p_1 increases. This suggests that the confidence set exhibits increased sensitivity to the outlier. Comparing within each column reveals that the confidence set also shifts towards the outlier as p_2 rises. To further quantify this effect, we conduct a sensitivity analysis of input vector estimation in Appendix C.2. These findings indicate a higher robustness against the influence of outliers in norms with lower values of p_1 and penalty functions with smaller values of p_2.

Asymmetric Confidence Set We now undertake the $\widehat{\text{MVCS}}$ approach for outcomes adhering to a skew (*i.e.*, asymmetric) distribution. For better visualization, we focus on a two-dimensional input vector $\tilde{\boldsymbol{z}} = (\tilde{z}_1, \tilde{z}_2)$ where $\tilde{z}_1 = \tilde{v}_1 + 0.2\tilde{v}_2$, $\tilde{z}_2 = \tilde{v}_2$, and $(\tilde{v}_1, \tilde{v}_2)$ are i.i.d. as a skew-normal distribution with a shape parameter $\alpha = 10$ (Azzalini and Capitanio 1999). Based on 100 samples, we solve $\widehat{\text{MVCS}}$ with the L_{p_1}-norm and a power penalty function $\psi(r) = r^{p_2}$. In particular, we apply the iterative method to Problem (5.13) to find a local optimum of Problem (5.12). The iterations repeat until the percentage reduction in the volume of the asymmetric confidence set falls below the threshold 0.002. We compare the asymmetric confidence set with a symmetric one of the same setting (*e.g.*, $\gamma = 1$), so that both have the same guarantee from the statistical principle as per Theorem 5.2. The results are summarized in Figure 5.2.

As expected, by taking skewness into account, $\widehat{\text{MVCS}}$ fits the skew data better. It can be observed that for all values of (p_1, p_2), to achieve the same statistical guarantee (*i.e.*, $\phi(r)$ in Theorem 5.2), the asymmetric confidence set always has a smaller volume compared with the

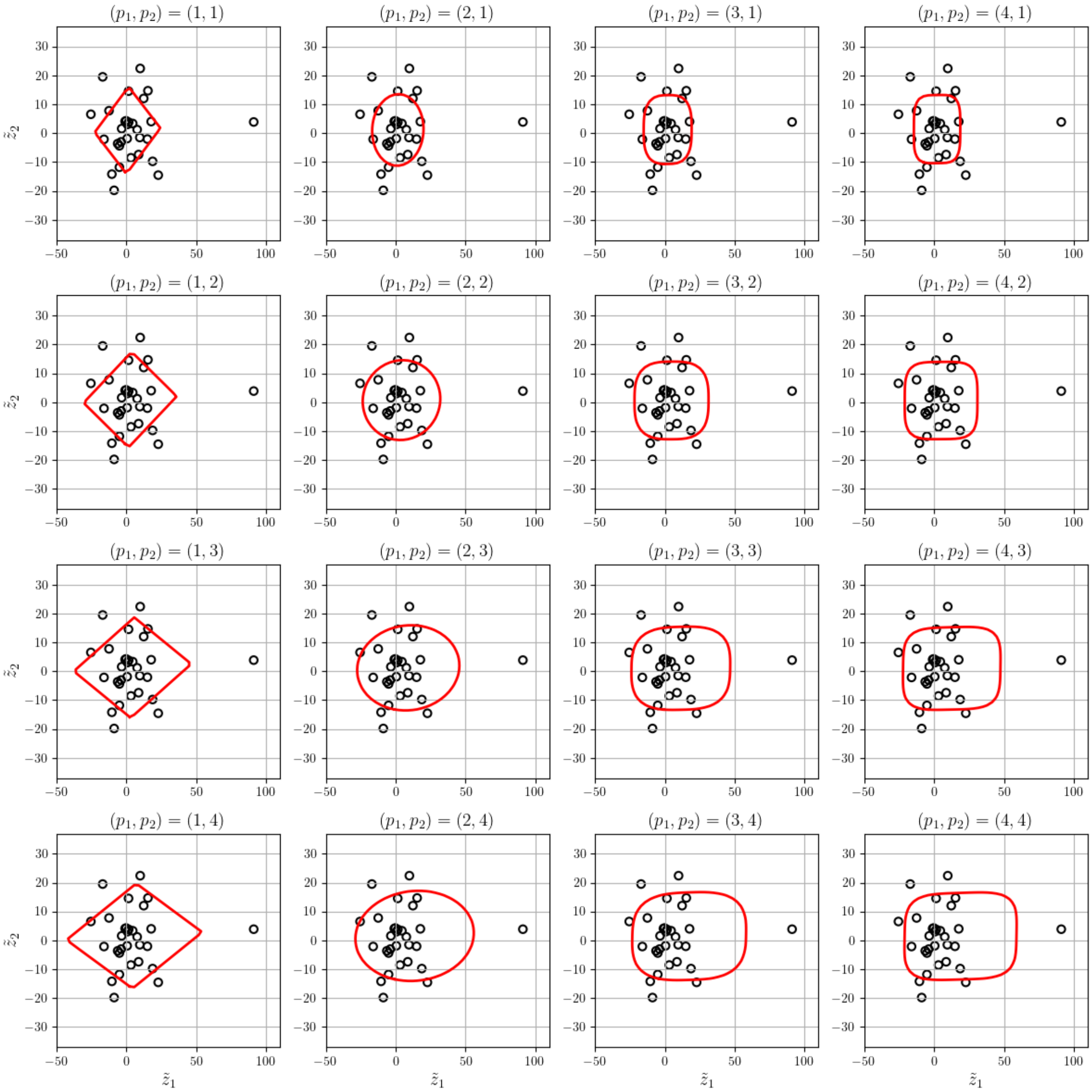


Figure 5.1: Symmetric $\widehat{\text{MVCS}}$ confidence sets with outliers.

Notes. In this replication, we generate 30 historical data points, and there is only one outlier along $\tilde{z}_1$'s dimension.

corresponding symmetric one. We also conduct an out-of-sample test based on 10000 samples of $\tilde{\boldsymbol{z}}$ to evaluate the probability that the realized input vector under the true distribution resides in the confidence sets constructed; see Figure 5.3 where we fix $p_1 = 2$ and $p_2 = 1$. The left panel shows that the asymmetric confidence set captures the probability density function more accurately. The right panel presents that, at a 95% confidence level, the asymmetric confidence set is 21% smaller in volume than the asymmetric one.

Asymmetric Confidence Set with Side Information Since the $\widehat{\text{MVCS}}$ approach supports asymmetric confidence sets as well as predictions with side information, we next consider an example concerning the linear relationships $\tilde{z}_1 = 3s - 2.5 + \tilde{\epsilon}_1$ and $\tilde{z}_2 = 5.5s - 10.5 + \tilde{\epsilon}_2$. The errors follow a standard skew-normal distribution with $\alpha = 10$, where ϵ_1 (resp., ϵ_2) is right-

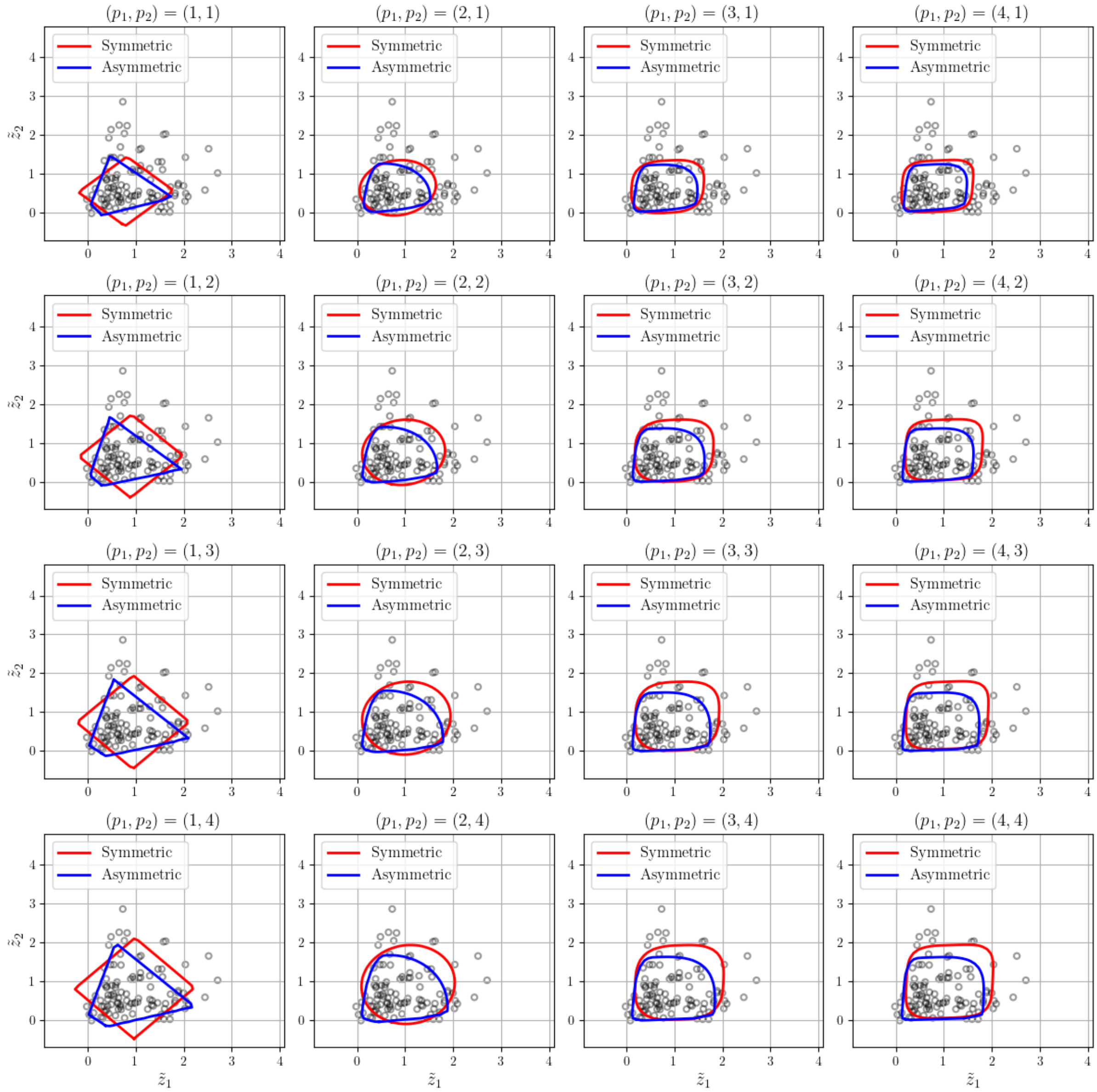


Figure 5.2: Symmetric and asymmetric $\widehat{\text{MVCS}}$ confidence sets for the skew-normal distribution.

skewed (resp., left-skewed). We plot the residuals and confidence sets in Figure 5.4. Comparing the contour maps, we observe that the asymmetric confidence set captures the skewness and fits the data more accurately. At the contour of value 2, the asymmetric confidence set includes all historical points whereas two points lie outside the symmetric confidence set.

5.6.2 Emergency service resource allocation

We examine an emergency service resource allocation problem on a network comprising L nodes. Each node $\ell \in [L]$ can store inventory to meet demands either at its own location or by transshipping to some neighboring nodes. The inventory at location $\ell \in [L]$ can only serve a specific subset of locations, denoted by $\mathcal{I}_\ell \subseteq [L]$. Likewise, the demand at location $\ell \in [L]$ can only be covered by the inventories transhipped from a subset of locations $\mathcal{J}_\ell \subseteq [L]$. In this exper-

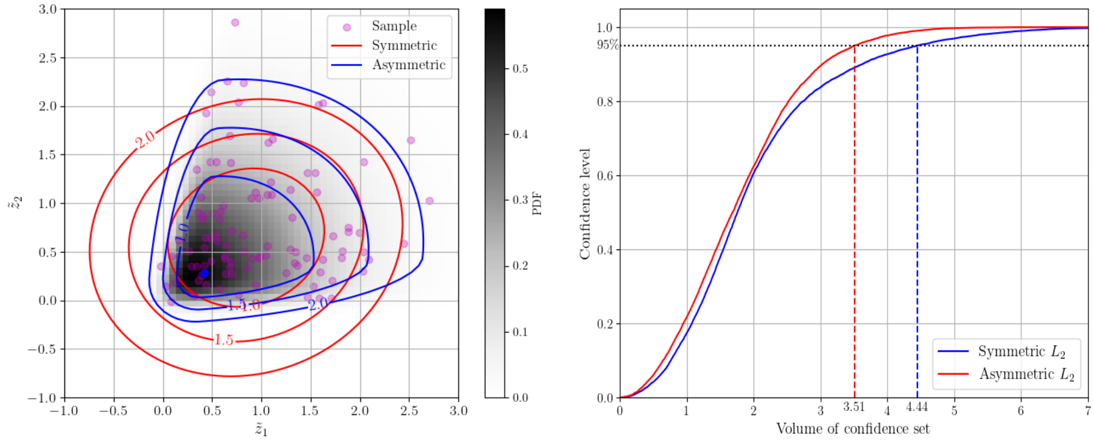


Figure 5.3: Confidence levels of symmetric and asymmetric confidence sets.

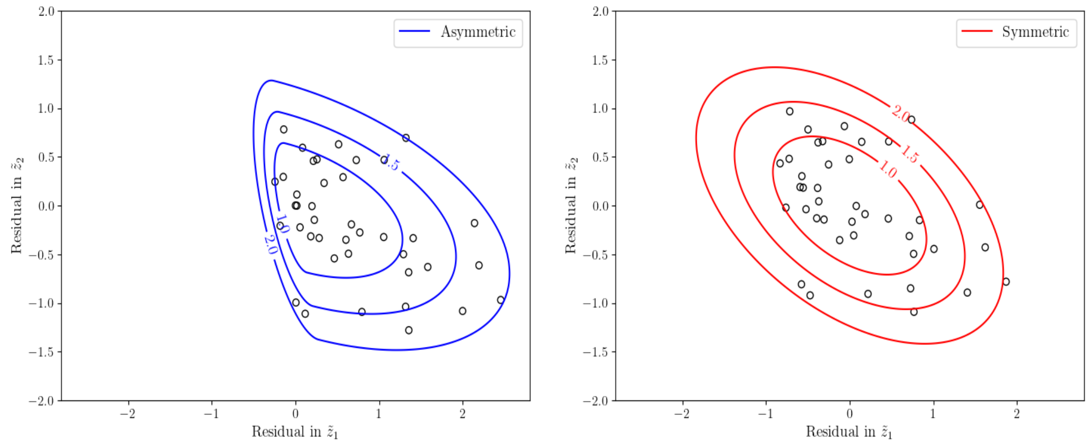


Figure 5.4: Symmetric and asymmetric $\widehat{\text{MVCS}}$ confidence sets with side information.

iment, we consider an undirected graph, therefore $\mathcal{J}_\ell = \mathcal{I}_\ell$. The uncertain parameters in this context are locational demands, collectively represented by the input vector $\boldsymbol{z} = (z_\ell)_{\ell \in [L]} \in \mathbb{R}^L$. The here-and-now decision vector $\boldsymbol{x} = (x_\ell)_{\ell \in [L]} \in \mathcal{X} \subseteq \mathbb{R}^L_+$ represents the inventory levels at locations. The wait-to-see decision y_{ij} denotes the transshipment quantity from location i to location j.

We define the cost-based attribute function that relates to the fulfillment shortfall as follows:

$$f(\boldsymbol{x}, \boldsymbol{z}) = \begin{cases} \min & s \\ \text{s.t.} & \displaystyle\sum_{j \in \mathcal{I}_\ell} y_{\ell j} \leq x_\ell \qquad \forall \ell \in [L] \\ & s \geq z_\ell - \displaystyle\sum_{i \in \mathcal{I}_\ell} y_{i\ell} \qquad \forall \ell \in [L] \\ & y_{\ell j} \geq 0 \qquad \forall j \in \mathcal{I}_\ell,\ \ell \in [L] \\ & s \in \mathbb{R}_+,\ \boldsymbol{x} \in \mathcal{X}. \end{cases} \tag{5.18}$$

When $f(\boldsymbol{x}, \boldsymbol{z}) > 0$, not all demands at the nodes can be met via transshipment, and a fulfillment shortfall occurs. One practical application of this problem involves the allocation of time-critical resources (such as ambulances and ventilators), which is widely known as the emergency medical service (EMS) problem (Boutilier and T. Chan 2020; Peng, et al. 2020). For the EMS problem, resources are constrained to meet local or neighboring demands, while operational costs (such as transportation expenses) are often disregarded due to their relatively low impact compared to procurement costs. Instead, coverage (*i.e.*, shortfall) is a more critical criterion, as a single shortage can result in severe consequences for patients. Therefore, our experiment will emphasize criteria such as the infeasible probability and the expected shortfall. We denote c_ℓ, $\ell \in [L]$ as the unit inventory cost at node ℓ. The historical demand samples are denoted by $\boldsymbol{z}_\omega \in \mathbb{R}^L$ for $\omega \in [\Omega]$.

To capture that demands should be fulfilled with high probability, we consider the scenario sampling approach (Calafiore and Campi, 2005) as our benchmark. This approach solves the following empirical optimization problem:

$$V_{\mathrm{EO}} = \left\{ \begin{array}{lll} \min & \boldsymbol{c}^\top \boldsymbol{x} & \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}_\omega) \le 0 & \forall \omega \in [\Omega] \\ & \boldsymbol{x} \in \mathcal{X} & \end{array} \right. = \left\{ \begin{array}{lll} \min & \boldsymbol{c}^\top \boldsymbol{x} & \\ \text{s.t.} & \sum\limits_{j \in \mathcal{I}_\ell} y_{\ell j \omega} \le x_i & \forall \ell \in [L],\ \omega \in [\Omega] \\ & 0 \ge z_{\ell\omega} - \sum\limits_{i \in \mathcal{I}_\ell} y_{i\ell\omega} & \forall \ell \in [L],\ \omega \in [\Omega] \\ & y_{\ell j \omega} \ge 0 & \forall j \in \mathcal{I}_\ell,\ \ell \in [L],\ \omega \in [\Omega] \\ & \boldsymbol{x} \in \mathcal{X}. & \end{array} \right. \tag{EO}$$

Calafiore and Campi (2005) show that if the samples are independently distributed, then with probability at least $1 - \frac{L}{(\Omega+1)\varepsilon}$, the solution to the EO problem remains feasible at $1 - \varepsilon \in [0, 1]$ confidence level. Consequently, achieving a higher level of feasibility may require a larger number of samples than are available.

We then use the optimal objective of EO as the reference for setting the targets for the robust decision-making models (*e.g.*, $V_0 = k_0 V_{\mathrm{EO}}$ for Step 0 of the framework in §5.1.2). Then, we compare the EO solution with the solutions obtained from the budgeted robust optimization problem, whose formulation is as follows:

$$\begin{array}{lll} \min & \boldsymbol{c}^\top \boldsymbol{x} & \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}) \le 0 & \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{x} \in \mathcal{X}. & \end{array}$$

Here, the uncertainty set is constructed as

$$\mathcal{Z}(\gamma) = \mathcal{Z} \cap \mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma) = \left\{ \boldsymbol{z} \in \mathcal{Z} \mid \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\| \le \gamma \right\},$$

where $\mathbf{D}$ and $\hat{\boldsymbol{z}}$ are estimated from solving the $\widehat{\text{MVCS}}$ problem and setting the outcome disparity metric as $\Delta(\boldsymbol{z}) = \|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|$. The budgeted robust optimization problem can be expressed as

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \sum_{j \in \mathcal{I}_\ell} y_{\ell j}(\boldsymbol{z}) \leq x_i && \forall \ell \in [L],\ \boldsymbol{z} \in \mathcal{Z}(\gamma) \\
& 0 \geq z_\ell - \sum_{i \in \mathcal{I}_\ell} y_{i\ell}(\boldsymbol{z}) && \forall \ell \in [L],\ \boldsymbol{z} \in \mathcal{Z}(\gamma) \\
& y_{\ell j}(\boldsymbol{z}) \geq 0 && \forall j \in \mathcal{I}_\ell,\ \ell \in [L],\ \boldsymbol{z} \in \mathcal{Z}(\gamma) \\
& y_{\ell j} \in \mathcal{R}(L) && \forall j \in \mathcal{I}_\ell,\ \ell \in [L] \\
& \boldsymbol{x} \in \mathcal{X},
\end{aligned}
\tag{5.19}
$$

where $\mathcal{R}(L)$ represents the set of all measurable functions on domain $\mathbb{R}^L$. The two-stage robust optimization problem (5.19) ensures that there is no shortfall even under the worst-case scenario in the uncertainty set. However, this two-stage problem is usually hard (if not impossible) to optimize, as the wait-to-see decisions are arbitrary functions of $\boldsymbol{z}$ (Ben-Tal, et al. 2004; Atamtürk and M. Zhang 2007). To obtain a tractable formulation, we use the affine recourse adaptation on the lifted uncertainty set advocated by Z. Chen, et al. (2023):

$$
\bar{\mathcal{Z}}(\gamma) = \left\{ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \mathcal{Z} \times \mathbb{R}^L_+ \times \mathbb{R}^L_+ \times \mathbb{R}_+ \;\middle|\; \begin{array}{l} \bar{\boldsymbol{w}} - \underline{\boldsymbol{w}} = \mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}) \\ \|\underline{\boldsymbol{w}} + \bar{\boldsymbol{w}}\| \leq u \leq \gamma \end{array} \right\}.
$$

Consequently, a tractable safe approximation of Problem (5.19) can be formulated as the following robust optimization problem

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \sum_{j \in \mathcal{I}_\ell} y_{\ell j}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \leq x_\ell && \forall \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}}(\gamma) \\
& 0 \geq z_\ell - \sum_{i \in \mathcal{I}_\ell} y_{i\ell}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) && \forall \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}}(\gamma) \\
& y_{\ell j}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \geq 0 && \forall j \in \mathcal{I}_\ell,\ \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}}(\gamma) \\
& y_{\ell j} \in \mathcal{A}(2L+1) && \forall j \in \mathcal{I}_\ell,\ \ell \in [L] \\
& \boldsymbol{x} \in \mathcal{X}.
\end{aligned}
\tag{RO}
$$

Here, $\mathcal{A}(2L+1)$ denotes the set of all affine functions on domain $\mathbb{R}^{2L+1}$. That is to say, a transshipment recourse function $y_{ij} : \mathbb{R}^{2L+1} \to \mathbb{R}$, as an affine function of $(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u)$ is given by:

$$
y_{ij}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) = y_{ij0} + \underline{\boldsymbol{y}}_{ij}^\top \underline{\boldsymbol{w}} + \bar{\boldsymbol{y}}_{ij}^\top \bar{\boldsymbol{w}} + y_{ij}^\star u
$$

for some $y_{ij0}, y_{ij}^\star \in \mathbb{R}$ and $\underline{\boldsymbol{y}}_{ij}, \bar{\boldsymbol{y}}_{ij} \in \mathbb{R}^L$. These are affine functions with respect to lifted variables.

We also consider the tolerated robust satisficing problem—Step 2 of the framework in §5.1.2 that specifies $\gamma^\star = 0$ and skips Step 1—whose approximation based on the affine recourse adaptation is formulated as follows:

$$
\begin{array}{lll}
\min & \kappa & \\
\text{s.t.} & \boldsymbol{c}^\top \boldsymbol{x} \leq \tau & \\
& \displaystyle\sum_{j \in \mathcal{I}_\ell} y_{\ell j}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \leq x_\ell & \forall \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}}(\bar{\gamma}) \\
& \kappa u \geq z_\ell - \displaystyle\sum_{i \in \mathcal{I}_\ell} y_{i\ell}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) & \forall \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}}(\bar{\gamma}) \\
& y_{\ell j}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \geq 0 & \forall j \in \mathcal{I}_\ell,\ \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}}(\bar{\gamma}) \\
& y_{\ell j} \in \mathcal{A}(2I+1) & \forall j \in \mathcal{I}_\ell,\ \ell \in [L] \\
& \boldsymbol{x} \in \mathcal{X}. &
\end{array}
\tag{RS}
$$

Here, $\bar{\gamma} \in \mathbb{R}_+$ is the maximum possible disparity, and we set it as the highest observed value in the training sample. The target τ can be chosen based on EO, for instance, $\tau = kV_{\mathrm{EO}}$ for some $k > 0$.

To make a fair comparison, for RO we use the bisection search to maximize γ such that its corresponding first-stage cost is identical to that of the RS Problem. This is equivalent to solving a budgeted satisficing problem—Step 1 of the framework in §5.1.2. By varying the target τ, we can generate a frontier for each model. These two robust optimization problems can be conveniently deployed by using the modeling toolkit, RSOME (Z. Chen, et al., 2020; Z. Chen and Xiong, 2023).

We randomly generate 15 nodes from a 10×10 square grid and construct the graph with 30 shortest arcs; the network is presented in Figure C.8 in Appendix C.3. Specifically, we assume that the unit inventory cost is proportional to the distance to the central point, that is, $c_i = 1 + d_i/5$. The maximum allocation level at each node is 200; therefore, $\mathcal{X} = [0, 200]^{15}$. The random demand $\tilde{z}_i$ follows a normal distribution modulated by an emergency factor. Specifically, $\tilde{z}_i \sim \mathbb{N}(\mu_i, \mu_i^2)\cdot(1+\tilde{v})$, where the rate parameter μ_i is randomly chosen from $\mathbb{U}[10, 30]$ and the term $\tilde{v}$ represents the emergent event, modeled as a Bernoulli random variable with a probability of occurrence of 0.05. Once the emergent event occurs, all demands are doubled. We then truncate each demand between 0 and 200, yielding $\mathcal{Z} = [0, 200]^{15}$. For evaluating the performance, we first generate $\Omega = 40$ samples to train $\widehat{\mathrm{MVCS}}$ and solve three problems. Additionally, we generate 2000 testing samples and recourse the second stage problem. We replicate the process 20 times to get the out-of-sample performance, evaluated at the following two performance metrics:

- *violation probability* as the percentage of scenarios that the demand can not be fulfilled;
- *average total violation* as the expected number of unfulfilled demands.

Note that we obtain 20 frontiers that correspond to 20 replications, among which we plot in Figure 5.5 the median and the 90% confidence interval (CI) between the 5th percentile and the

95th percentile of the infeasibility profile in Figure 5.5. The performance of three models at the same level of first-stage cost in each replication is provided in Appendix C.3.

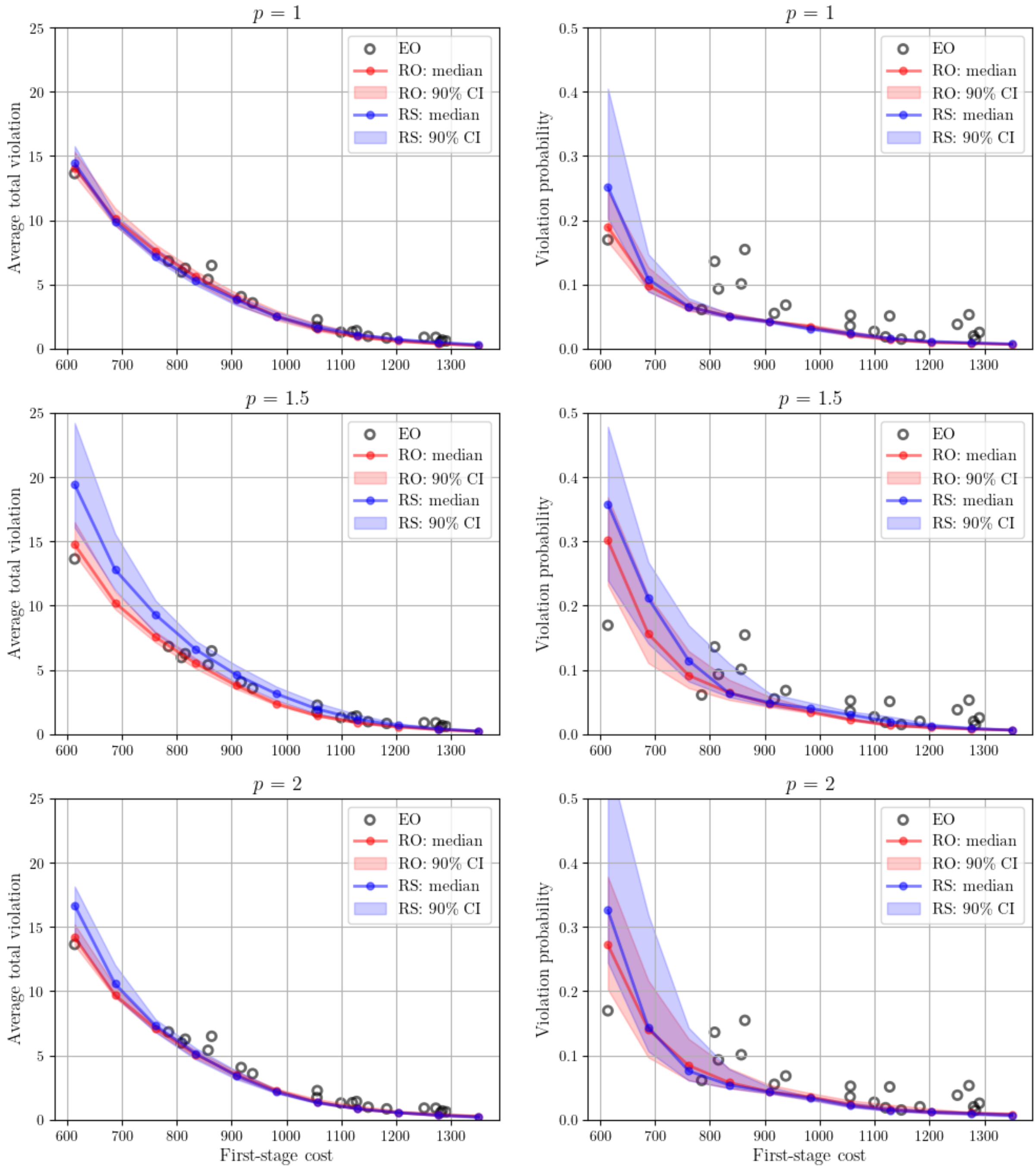


Figure 5.5: Infeasibility profile of EO, RO and RS models with $\Omega = 40$ training samples.

It can be seen that both RO and RS models with L_1- and L_2-norms outperform the EO model in both performance metrics. The models with $L_{1.5}$-norm show better performance than EO in view of violation probability. Such an improved performance can be attributed to the EO model's heavy reliance on the training samples, resulting in unstable performance, particularly in the other performance metric of violation probability. Although EO can avoid infeasibility in 40 training samples, it could provide less protection in the out-of-sample testing. Furthermore, both robust decision-making frameworks can achieve lower violation probabilities by increasing the target value, and they demonstrate comparable performance at high target values, with RS with L_2-norm performing slightly better. Specifically, at a cost level of 1055.50, RS with L_2-norm has an average total violation of 1.33, whereas RS with L_1-norm has an average total violation of 1.67 and RO with L_2-norm has an average total violation of 1.37. They all have violation probabilities around 0.02. Additionally, RO has more stable performance compared to RS at low target values, such as 614.15 and 687.71. This finding underscores the necessity of conducting Step 1 before Step 2 in our framework in §5.1.2.

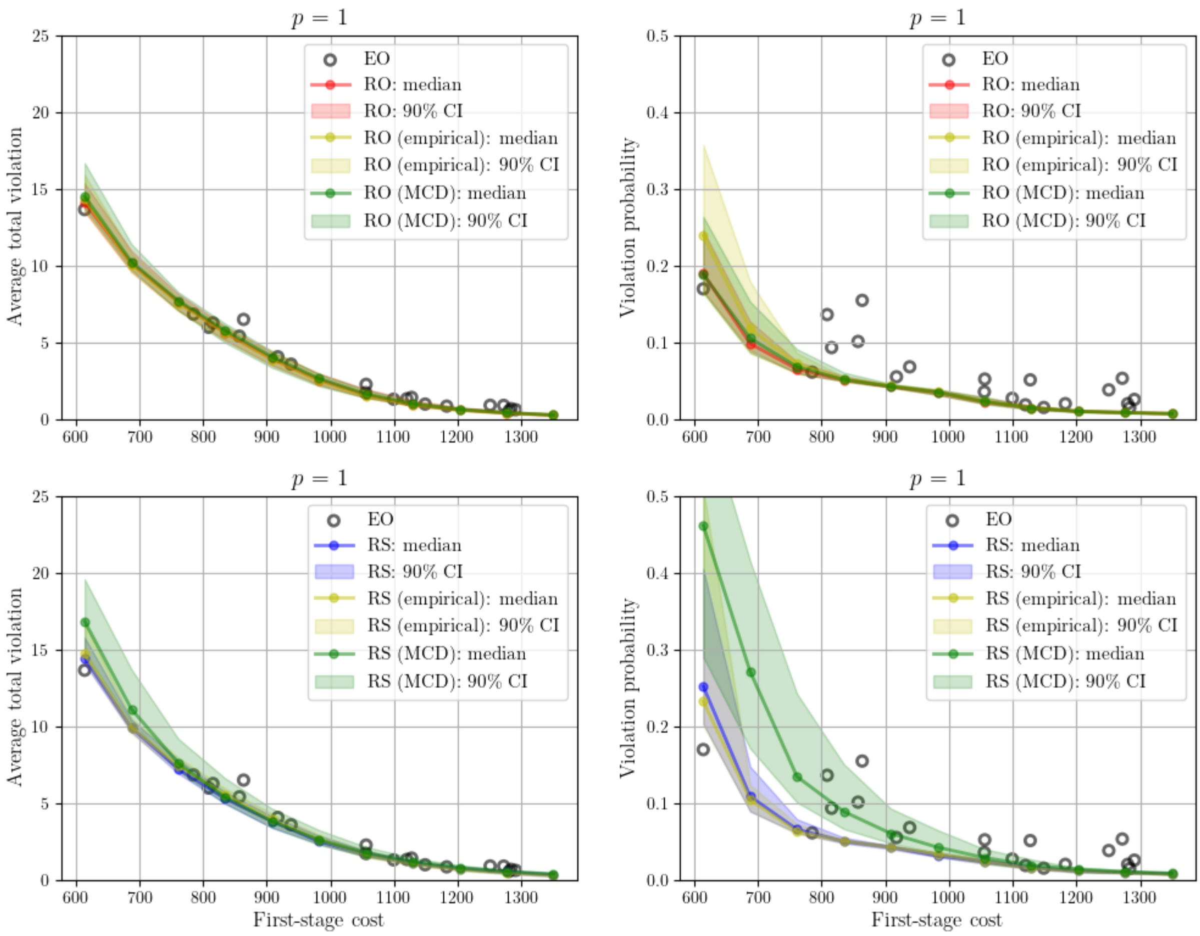


Figure 5.6: Infeasible profile of 3 models without outliers ($\Omega = 40$ training samples).

Finally, we compare outcomes of RO and RS models, employing distinct estimation methods for the parameters $\hat{z}$ and $\mathbf{D}$. These methods include the sample mean and covariance (empirical) and MCD. Our investigation focuses on robust decision-making problems formulated with the L_1-norm. Following the same analysis procedure, we evaluate the out-of-sample infeasibility profiles. Results derived from uncontaminated training samples are presented in Figure 5.6, while those from contaminated training samples are illustrated in Figure 5.7. To simulate contamination, 5% of the training samples are modified by replacing the demand data for a fixed single node with the upper bound value of 200.

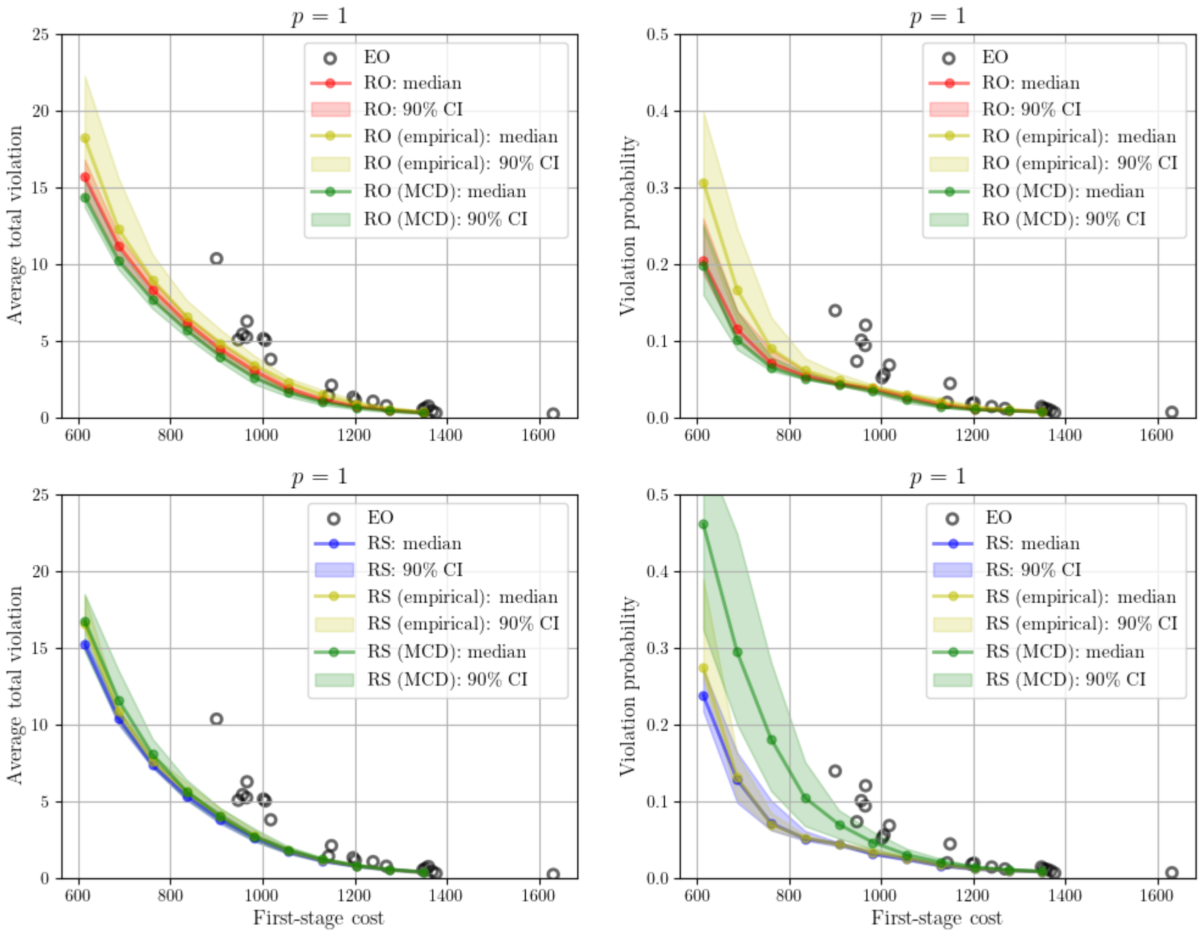


Figure 5.7: Infeasible profile of 3 models with outliers ($\Omega = 40$ training samples).

Notably, the performance of EO exhibits significant degradation under these contaminated conditions. It allocates excess inventory to nodes that fail to reduce violation risks, highlighting its pronounced dependence on high-quality data. When comparing varying estimations of $\hat{\boldsymbol{z}}$ and $\mathbf{D}$, robust decision-making models using the $\widehat{\text{MVCS}}$ method demonstrate the most stable performance across conditions. In contrast, empirical estimators yield suboptimal results in RO, while the MCD approach demonstrates significant instability in RS. These findings highlight the reliability of the $\widehat{\text{MVCS}}$ method.

5.7 Conclusion

We propose a general framework with comprehensive robust constraints that addresses two layers of uncertainty—outcome uncertainty and estimation uncertainty. This framework synergizes the prevalent robust optimization techniques with the emerging robust satisficing paradigm. As Alfred Whitehead stated, "The art of progress is to preserve order amid change and to preserve change amid order." In our unifying sequential robust satisficing decision framework, the "hard" constraints in *the budgeted robust satisficing* phase (Steps 1 and 3) enhance the stability, and the "soft" constraints in *the tolerated robust satisficing* phase (Steps 1 and 3) allow some degree of flexibility in input vectors. By these phases, our approach aims to reduce the conservativeness of robust solutions while enhancing their feasibility. Moreover, the combination of outcome and estimation uncertainties can provide better protection and get more stable performance, as illustrated in the portfolio optimization example. We introduce a novel MVCS approach to determine parameters in *outcome disparity metrics*. The $\widehat{\text{MVCS}}$ approach is developed with expected deviation penalty functions over empirical distributions, and this data-driven method has demonstrated adaptability in learning various patterns from data, including skewness and side information. In specific settings, it shows connections to traditional estimation approaches. Additionally, we illustrate that different norms and disparity penalty functions exhibit varying levels of robustness against outliers. In summary, our contributions enhance methodologies that learn uncertainty from data and achieve a more favorable trade-off against the *price of robustness*.

Appendix

C.1 Proofs

Proof Theorem 5.1

Proof. Fixing $\boldsymbol{z} \in \mathcal{Z}$, for all $r \geq 0$ and $\boldsymbol{\zeta} \in \hat{\mathcal{Z}}$, any $\boldsymbol{x}$ feasible to the comprehensive robust constraint (5.3) satisfies

$$f(\boldsymbol{x}, \boldsymbol{z}) - \kappa(\Delta(\boldsymbol{z}, \boldsymbol{\zeta}) - \gamma)^+ - \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+ \leq \tau.$$

Hence, for all $r \geq 0$ and $\boldsymbol{\zeta} \in \mathcal{Z}$, we can deduce that

$$\begin{aligned}
& \mathbb{P}\left\{f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r + \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+\right\} \\
\leq\ & \mathbb{P}\big\{f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \kappa(\Delta(\tilde{\boldsymbol{z}}, \boldsymbol{\zeta}) - \gamma)^+ - \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+ + \kappa r + \hat{\kappa}(\hat{\Delta}(\boldsymbol{\zeta}, \hat{\boldsymbol{z}}) - \hat{\gamma})^+\big\} \\
=\ & \mathbb{P}\left\{\kappa(\Delta(\tilde{\boldsymbol{z}}, \boldsymbol{\zeta}) - \gamma)^+ > \kappa r\right\} \\
\leq\ & \mathbb{P}\left\{\Delta(\tilde{\boldsymbol{z}}, \boldsymbol{\zeta}) > r + \gamma\right\} \\
=\ & \mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\boldsymbol{\zeta}, r + \gamma)\right\},
\end{aligned}$$

where we note that the second inequality indeed holds for any $\kappa \in [0, \infty]$.

□

Proof of Theorem 5.2

Proof. For any $\gamma \geq 0$, we have

$$\begin{aligned}
\mathbb{P}\left\{\tilde{\boldsymbol{z}} \notin \mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)\right\} &= \mathbb{P}\left\{\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\| > \gamma\right\} \\
&\leq \mathbb{P}\left\{\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|) > \psi(\gamma)\right\} \\
&\leq \frac{\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right]}{\psi(\gamma)} \\
&\leq \frac{1}{\psi(\gamma)} \\
&\leq \varphi(\gamma),
\end{aligned}$$

where the first inequality is because the deviation penalty function ψ is non-decreasing in $\gamma \geq 0$ and the second inequality applies Markov inequality as ψ is non-negative.

$\square$

Proof of Theorem 5.3

Proof. The MVCS problem is equivalent to

$$\begin{array}{ll} \inf & \det(\mathbf{D}) \\ \text{s.t.} & \min_{\hat{\boldsymbol{z}} \in \mathbb{R}^L} \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \leq 1 \\ & \mathbf{D} \in \mathbb{S}^L_{++}. \end{array} \tag{C.20}$$

Observe that the function $g(\hat{\boldsymbol{z}}) = \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right]$ is convex in $\hat{\boldsymbol{z}}$. It thus suffices to show that

$$g(\boldsymbol{\mu} + \boldsymbol{d}) \geq g(\boldsymbol{\mu}) \quad \forall \boldsymbol{d} \in \mathbb{R}^L.$$

Indeed, by the convexity of g, we have for all $\boldsymbol{d} \in \mathbb{R}^L$,

$$\begin{array}{rcl} 2g(\boldsymbol{\mu}) & \leq & g(\boldsymbol{\mu} + \boldsymbol{d}) + g(\boldsymbol{\mu} - \boldsymbol{d}) \\ & = & g(\boldsymbol{\mu} + \boldsymbol{d}) + \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \boldsymbol{\mu} + \boldsymbol{d})\|)\right] \\ & = & g(\boldsymbol{\mu} + \boldsymbol{d}) + \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(-\tilde{\boldsymbol{z}} + \boldsymbol{\mu} + \boldsymbol{d})\|)\right] \\ & = & g(\boldsymbol{\mu} + \boldsymbol{d}) + \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \boldsymbol{\mu} - \boldsymbol{d})\|)\right] \\ & = & g(\boldsymbol{\mu} + \boldsymbol{d}) + g(\boldsymbol{\mu} + \boldsymbol{d}) \\ & = & 2g(\boldsymbol{\mu} + \boldsymbol{d}). \end{array}$$

The second equability is because the distribution is symmetrical about the mean and the third property is due to the absolute homogeneity of norms. Hence, $\hat{\boldsymbol{z}} = \boldsymbol{\mu}$ is optimal.

$\square$

Proof of Theorem 5.4

Proof. Note that we can express

$$\mathbb{E}_{\mathbb{P}}\left[\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_2^2\right] = \mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})^\top \mathbf{D}^{-2}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\right].$$

With the change of decision variable from $\mathbf{D}^{-2}$ to $\mathbf{W}$, and noting that

$$\log\det(\mathbf{D}) = \log\det(\mathbf{W}^{-1/2}) = -\frac{1}{2}\log\det(\mathbf{W}),$$

we can express Problem (5.6) equivalently as the following optimization problem:

$$\begin{array}{ll} \sup & \log\det(\mathbf{W}) \\ \text{s.t.} & \mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}})^\top \mathbf{W}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}))\right] \le L \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \mathbf{W} \in \mathbb{S}^L_{++}. \end{array} \tag{C.21}$$

Fix any $\mathbf{W} \in \mathbb{S}^L_{++}$ and consider the following convex quadratic function

$$g(\hat{\boldsymbol{z}}) = \mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}})^\top \mathbf{W}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}))\right] = \hat{\boldsymbol{z}}^\top \mathbf{W}\hat{\boldsymbol{z}} - 2\boldsymbol{\mu}^\top \mathbf{W}\hat{\boldsymbol{z}} + \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{z}}^\top \mathbf{W}\tilde{\boldsymbol{z}}\right],$$

which has an unique minimum at $\hat{\boldsymbol{z}}^\star = \boldsymbol{\mu}$. Crucially, this optimum is independent of $\mathbf{W}$. Hence, $\hat{\boldsymbol{z}}^\star = \boldsymbol{\mu}$ is also part of an optimal solution to Problem (C.21). Using this solution, we can represent

$$\mathbb{E}_{\mathbb{P}}\left[(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})^\top \mathbf{W}(\tilde{\boldsymbol{z}}-\boldsymbol{\mu}))\right] = \mathbb{E}_{\mathbb{P}}\left[\operatorname{tr}\left(\mathbf{W}(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})^\top\right)\right] = \operatorname{tr}(\mathbf{W}\boldsymbol{\Sigma}),$$

where the second equation follows from the fact that the trace operator is linear, allowing us to interchange the expectation and the trace. Hence, Problem (C.21) can be simplified as follows:

$$\begin{array}{ll} \sup & \log\det(\mathbf{W}) \\ \text{s.t.} & \operatorname{tr}(\mathbf{W}\boldsymbol{\Sigma}) \le L \\ & \mathbf{W} \in \mathbb{S}^L_{++}. \end{array} \tag{C.22}$$

The objective function of Problem (C.22) is concave in $\mathbf{W}$ and its Lagrangian function is

$$\Lambda(\mathbf{W}, \lambda) = \log\det(\mathbf{W}) - \lambda\left(\operatorname{tr}(\mathbf{W}\boldsymbol{\Sigma}) - L\right)$$

for some $\lambda \ge 0$. Taking the derivative and equating to zero, the optimal solutions $\mathbf{W}^\star$ and $\lambda^\star$ satisfy:

$$\mathbf{W}^\star = \frac{\boldsymbol{\Sigma}^{-1}}{\lambda^\star} \quad \text{and} \quad \operatorname{tr}(\mathbf{W}^\star\boldsymbol{\Sigma}) = L.$$

Hence, $\operatorname{tr}(\mathbf{I})/\lambda^\star = L$, implying $\lambda^\star = 1$ and $\mathbf{W}^\star = \boldsymbol{\Sigma}^{-1}$. Therefore, $\mathbf{D}^\star = \boldsymbol{\Sigma}^{1/2}$.

□

Proof of Theorem 5.5

Proof. Since Gaussian distribution is symmetric, we have from Theorem 5.3 that $\hat{\boldsymbol{z}}^\star = \boldsymbol{\mu} = \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}]$. Hence, with the change of decision variable from $\mathbf{D}^{-1}$ to $\mathbf{Q}$, Problem (C.20) is equivalent to

$$
\begin{array}{ll}
\sup & \log\det(\mathbf{Q}) \\
\text{s.t.} & \mathbb{E}_{\mathbb{P}}\left[\|\mathbf{Q}(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})\|_p^p\right] \leq L \\
& \mathbf{Q} \in \mathbb{S}_{++}^L.
\end{array} \tag{C.23}
$$

Suppose $\mathbf{Q} = [\boldsymbol{q}_1 \ \cdots \ \boldsymbol{q}_L]^\top$, under Gaussian distribution we have

$$
\mathbb{E}_{\mathbb{P}}\left[\|\mathbf{Q}(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})\|_p^p\right] = \mathbb{E}_{\mathbb{P}}\left[\sum_{\ell\in[L]} |\boldsymbol{q}_\ell^\top(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})|^p\right] = \sum_{\ell\in[L]} \mathbb{E}_{\mathbb{P}}\left[|\boldsymbol{q}_\ell^\top(\tilde{\boldsymbol{z}}-\boldsymbol{\mu})|^p\right] = \sum_{\ell\in[L]} (\mathbf{q}_\ell^\top \boldsymbol{\Sigma} \mathbf{q}_\ell)^{p/2}\beta,
$$

where $\boldsymbol{q}_\ell^\top(\tilde{\boldsymbol{z}}-\boldsymbol{\mu}) \sim \mathcal{N}(0, \boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)$ and

$$
\beta = \frac{2^{p/2}}{\sqrt{\pi}} \int_0^{+\infty} t^{\frac{p-1}{2}} e^{-t} \mathrm{d}t = \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right).
$$

In the last equation, we use the central absolute moments of Gaussian distribution such that

$$
\begin{aligned}
\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi \boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell}} |x|^p e^{-\frac{1}{2}x^2/(\boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)} \mathrm{d}x &= (\boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)^{p/2} \cdot 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} x^p e^{-\frac{1}{2}x^2} \mathrm{d}x \\
&= (\boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)^{p/2} \cdot \frac{2^{p/2}}{\sqrt{\pi}} \int_0^{+\infty} t^{\frac{p-1}{2}} e^{-t} \mathrm{d}t \\
&= (\boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)^{p/2} \beta.
\end{aligned}
$$

The Lagrangian function for Problem (C.23) is as follows:

$$
\Lambda(\mathbf{Q}, \lambda) = \log\det(\mathbf{Q}) - \lambda \sum_{\ell\in[L]} (\boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)^{p/2}\beta + \lambda L.
$$

Using Jacobi's formula, $\nabla_{\mathbf{Q}} \log\det(\mathbf{Q}) = \mathbf{Q}^{-1}$. We can then calculate the gradient with respect to $\boldsymbol{q}_\ell$ as follows:

$$
\nabla_{\boldsymbol{q}_\ell} \Lambda(\mathbf{Q}, \lambda) = \mathbf{Q}^{-1}\boldsymbol{e}_\ell - \lambda p \beta (\boldsymbol{q}_\ell^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell)^{\frac{p}{2}-1} \boldsymbol{\Sigma} \boldsymbol{q}_\ell. \tag{C.24}
$$

To verify the optimality of the solution $(\lambda^\star, \mathbf{Q}^\star) = (1/p, \beta^{-\frac{1}{p}} \boldsymbol{\Sigma}^{-\frac{1}{2}})$, we note that

$$
(\boldsymbol{q}_\ell^\star)^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell^\star = \beta^{-\frac{2}{p}} (\boldsymbol{e}_\ell^\top \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{e}_\ell) = \beta^{-\frac{2}{p}}.
$$

Substituting it into Equation (C.24), we have

$$
\nabla_{\boldsymbol{q}_\ell} \Lambda(\mathbf{Q}^\star, \lambda^\star) = \beta^{\frac{1}{p}} \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{e}_\ell - \lambda^\star p \beta \cdot \beta^{-\frac{2}{p}\cdot(\frac{p}{2}-1)} \cdot \beta^{-\frac{1}{p}} \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{e}_\ell = \beta^{\frac{1}{p}} \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{e}_\ell - \lambda^\star p \beta^{\frac{1}{p}} \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{e}_\ell = \mathbf{0}.
$$

Additionally, this solution satisfies

$$\sum_{\ell \in [L]} ((\boldsymbol{q}_\ell^\star)^\top \boldsymbol{\Sigma} \boldsymbol{q}_\ell^\star)^{\frac{p}{2}} \beta = \sum_{\ell \in [L]} \beta^{-1} \beta = L,$$

concluding its optimality since Problem (C.23) is a convex optimization problem.

□

Proof of Theorem 5.6

Proof. Suppose $\boldsymbol{z}$ is feasible to the set as defined in (5.7). Then it would also be feasible to the set as defined (5.8) for the lifted variables

$$\underline{\boldsymbol{w}} = (\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^- \in \mathbb{R}^L_+ \quad and \quad \bar{\boldsymbol{w}} = (\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^+ \in \mathbb{R}^L_+.$$

Conversely, we use the property of absolute norm derived in Bertsimas and Sim (2006) that for any $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{R}^L$ such that $v_\ell \geq |u_\ell|$ for all $\ell \in [L]$, $\|\boldsymbol{v}\| \geq \|\boldsymbol{u}\|$. Hence, if $\boldsymbol{z}$ is feasible in the set as defined in (5.8), then

$$\begin{aligned} & \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^+\| \\ = \; & \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})(\bar{\boldsymbol{w}} - \underline{\boldsymbol{w}})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})(\bar{\boldsymbol{w}} - \underline{\boldsymbol{w}})^+\| \\ \leq \; & \|\,\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\underline{\boldsymbol{w}} + \mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\bar{\boldsymbol{w}}\| \leq \gamma. \end{aligned}$$

This is because since $\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}} \geq 0$, we have $(\bar{\boldsymbol{w}} - \underline{\boldsymbol{w}})^- \leq \underline{\boldsymbol{w}}$ and $(\bar{\boldsymbol{w}} - \underline{\boldsymbol{w}})^+ \leq \bar{\boldsymbol{w}}$. Hence, $\boldsymbol{z}$ is also feasible to the set as defined in (5.7).

□

Proof of Theorem 5.7

Proof. Use the property of absolute norm, observe that

$$\begin{aligned}
&\int_{\{\boldsymbol{u}\in\mathbb{R}^L \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
= &\int_{\{\boldsymbol{u}\in\mathbb{R}_-\times\mathbb{R}^{L-1} \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
&+ \int_{\{\boldsymbol{u}\in\mathbb{R}_+\times\mathbb{R}^{L-1} \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
= &\ \underline{\sigma}_1 \int_{\{\boldsymbol{u}\in\mathbb{R}_-\times\mathbb{R}^{L-1} \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
&+ \bar{\sigma}_1 \int_{\{\boldsymbol{u}\in\mathbb{R}_+\times\mathbb{R}^{L-1} \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
= &\ (\underline{\sigma}_1 + \bar{\sigma}_1) \int_{\{\boldsymbol{u}\in\mathbb{R}_+\times\mathbb{R}^{L-1} \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
= &\ \frac{\underline{\sigma}_1 + \bar{\sigma}_1}{2} \int_{\{\boldsymbol{u}\in\times\mathbb{R}^{L} \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}}^\dagger)\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u},
\end{aligned}$$

where $\underline{\boldsymbol{\sigma}}^\dagger = (1, \underline{\sigma}_2, \ldots, \underline{\sigma}_L)$ and $\bar{\boldsymbol{\sigma}}^\dagger = (1, \bar{\sigma}_2, \ldots, \bar{\sigma}_L)$. We continue this transformation until both asymmetric normalization vectors become vectors of all ones. Hence, we have

$$\begin{aligned}
\mathrm{Vol}(\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma)) &= \int_{\{\boldsymbol{u}\in\mathbb{R}^L \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u} \\
&= \int_{\{\boldsymbol{u}\in\mathbb{R}^L \ | \ \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\boldsymbol{u})^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\boldsymbol{u})^+\|\le\gamma\}} \mathrm{d}\boldsymbol{u}\, \det(\mathbf{D}) \\
&= \prod_{\ell\in[L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \int_{\{\boldsymbol{u}\in\mathbb{R}^L \ | \ \|\boldsymbol{u}\|\le\gamma\}} \mathrm{d}\boldsymbol{u}\, \det(\mathbf{D}) \\
&= \prod_{\ell\in[L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \cdot \mathrm{Vol}(\mathcal{B}(\mathbf{0}, \mathbf{I}, \mathbf{1}, \mathbf{1}, \gamma)) \\
&\propto \prod_{\ell\in[L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}),
\end{aligned}$$

concluding the proof.

□

Proof of Theorem 5.8

Proof. Given the continuity of the deviation penalty function ψ at $r = 0$ and $\psi(0) < 1$, we note that the solution $(\boldsymbol{q}, \mathbf{Q}) = (\mathbf{0}, \epsilon\mathbf{I})$ is feasible for some $\epsilon > 0$. Therefore, the optimal solution must satisfy $\mathbf{Q} \in \mathbb{S}^L_{++}$. Consider the variable substitution from $\mathbf{D}^{-1}$ to $\mathbf{Q}$, which is possible because $\mathbf{D} \succ \mathbf{0}$ is a positive definite matrix. Note that $\mathbf{D} \in \mathbb{S}^L_{++}$ is equivalent to $\mathbf{Q} = \mathbf{D}^{-1} \in \mathbb{S}^L_{++}$. With another variable substitution $\hat{\boldsymbol{z}} = \mathbf{D}\boldsymbol{q}$, we then have the following equivalence:

$$\mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \le 1 \quad \Longleftrightarrow \quad \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|)\right] \le 1.$$

Given that ψ is a convex and non-increasing function, $\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|)\right]$ is a convex function with respect to the decision variable $(\boldsymbol{q}, \mathbf{Q})$. Hence, $\mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|)\right] \leq 1$ is a convex constraint.

Consider the case of having a positive regularization vector $\boldsymbol{r} > \mathbf{0}$. By Schur complement,

$$\mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r}) \quad \Longleftrightarrow \quad \mathbf{D} - \mathrm{Diag}(\boldsymbol{r}) \succeq \mathbf{0} \quad \Longleftrightarrow \quad \begin{bmatrix} \mathbf{D} & \mathbf{I} \\ \mathbf{I} & \mathrm{Diag}^{-1}(\boldsymbol{r}) \end{bmatrix} \succeq 0.$$

By left multiplying $\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{D}^{-1} & \mathbf{I} \end{bmatrix}$ and right multiplying its transpose, we obtain that

$$\begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathrm{Diag}^{-1}(\boldsymbol{r}) - \mathbf{D}^{-1} \end{bmatrix} \succeq 0,$$

which implies that $\mathrm{Diag}^{-1}(\boldsymbol{r}) \succeq \mathbf{D}^{-1} = \mathbf{Q}$. Since $\mathbf{Q}$ is also a positive definite matrix in Problem (5.9), by the same argument, the constraint $\mathrm{Diag}^{-1}(\boldsymbol{r}) \succeq \mathbf{Q}$ implies $\mathrm{Diag}(\boldsymbol{r}) \preceq \mathbf{Q}^{-1} = \mathbf{D}$. That is to say, the convex constraint $\mathbf{Q} \preceq \mathrm{Diag}^{-1}(\boldsymbol{r})$ after the variable substitution is equivalent to the original constraint $\mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r})$.

The objective of Problem ($\widehat{\mathrm{MVCS}}$) that minimizes $\det(\mathbf{D}) = 1/\det(\mathbf{Q})$, is equivalent to maximizing $\log\det(\mathbf{Q})$, for which the log determinant function is a concave function. This concludes that Problem (5.9) is a convex optimization problem.

□

Proof of Theorem 5.9

Proof. Consider the change of decision variable from d_i to q_i^{-1} for each $i \in [L]$, Problem (5.11) is then equivalent to

$$\begin{array}{ll} \max & \sum\limits_{i\in[L]} \log(q_i) \\ \text{s.t.} & \sum\limits_{i\in[L]} q_i^p \mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i|^p\right] \leq L \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \boldsymbol{q} \in \mathbb{R}^L_+. \end{array}$$

Observe that for each $i \in [L]$, the solution $\hat{z}_i^\star$ minimizes $\mathbb{E}_{\mathbb{P}}\left[|\tilde{z}_i - \hat{z}_i|^p\right]$. Since this is a convex function with respect to $\hat{z}_i$, the optimal solution occurs when there exists a zero subgradient. To arrive at the desired result, we note that for $p > 1$ the gradient of the function $|r|^p$ is $((r)^+)^{p-1} - ((r)^-)^{p-1}$, while for $p = 1$, the gradient of $|r|$ is $sign(r)$. Hence, we focus on the following optimization problem:

$$\begin{aligned}\max \quad & \sum_{i\in[L]} \log(q_i) \\ \text{s.t.} \quad & \sum_{i\in[L]} q_i^p \mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i^\star|^p\right] \le L \\ & \boldsymbol{q} \in \mathbb{R}_+^L.\end{aligned}$$

Solving this convex optimization problem by equating the gradient of its Lagrangian function to zero gives

$$\frac{1}{q_i^\star} = \lambda^\star p q_i^{\star p-1} \mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i^\star|^p\right]$$

and

$$\sum_{\ell\in[L]} q_i^{\star p} \mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i^\star|^p\right] = L.$$

This then yields $\lambda^\star = 1/p$ and $q_i^\star = \left[\mathbb{E}_{\hat{\mathbb{P}}_i}\left[|\tilde{z}_i - \hat{z}_i^\star|^p\right]\right]^{-1/p}$ for each $i \in [L]$.

□

Proof of Theorem 5.10

Proof. With the change of decision variables from $\underline{\sigma}_\ell$ and $\bar{\sigma}_\ell$ to their reciprocals $\underline{v}_\ell = \underline{\sigma}_\ell^{-1}$ and $\bar{v}_\ell = \bar{\sigma}_\ell^{-1}$ for each $\ell \in [L]$, the objective function of Problem (5.13) minimizes $\prod_{\ell\in[L]}\left(\frac{1}{2\underline{v}_\ell} + \frac{1}{2\bar{v}_\ell}\right)$. By introducing the variables, $\boldsymbol{t} \in \mathbb{R}^L$, we can reformulated Problem (5.13) to the following problem:

$$\begin{aligned}\sup \quad & \sqrt[L]{\textstyle\prod_{\ell\in[L]} t_\ell} \\ \text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|(\mathrm{Diag}(\underline{\boldsymbol{v}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}(\bar{\boldsymbol{v}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^+\|)\right] \le 1 \\ & \frac{1}{2\underline{v}_\ell} + \frac{1}{2\bar{v}_\ell} \le \frac{1}{t_\ell} & \forall \ell \in [L] \\ & \boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}_{++}^L.\end{aligned}$$

The second constraint, $\frac{1}{2\underline{v}_\ell} + \frac{1}{2\bar{v}_\ell} \le \frac{1}{t_\ell}$, can be expressed as the following set of constraints:

$$\frac{1}{2\underline{v}_\ell} \le \frac{\underline{w}_\ell}{t_\ell^2}, \quad \frac{1}{2\bar{v}_\ell} \le \frac{\bar{w}_\ell}{t_\ell^2}, \quad \frac{\underline{w}_\ell}{t_\ell^2} + \frac{\bar{w}_\ell}{t_\ell^2} \le \frac{1}{t_\ell},$$

for some $\underline{w}_\ell, \bar{w}_\ell \ge 0$. This then yields rotated second-order conic constraints as expressed in the Problem (5.14), concluding the proof.

□

Proof. of Theorem 5.11

Proof. Since the comprehensive robust constraint (5.3) can be expressed as

$$\sup_{\boldsymbol{z}\in\mathcal{Z}} \{f(\boldsymbol{x},\boldsymbol{z}) - \Phi(\boldsymbol{z})\} \leq \tau,$$

where

$$\Phi(\boldsymbol{z}) = \inf_{\boldsymbol{\zeta}\in\hat{\mathcal{Z}}} \{\kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) - \gamma)^+ + \hat{\kappa}(\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \hat{\gamma})^+\}.$$

It thus suffices to show

$$\Phi(\boldsymbol{z}) = \min\{\kappa,\hat{\kappa}\}(\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) - \gamma - \hat{\gamma})^+.$$

Since $\Phi(\boldsymbol{z}) \geq 0$ and $\Phi(\hat{\boldsymbol{z}}) = 0$ (achieved when $\boldsymbol{\zeta} = \hat{\boldsymbol{z}}$), we focus on $\boldsymbol{z} \neq \hat{\boldsymbol{z}}$, or equivalently, $\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) > 0$.

Suppose first $\hat{\kappa} \geq \kappa$. From the property of subjectivity associated with the disparity metric Δ and the function $(\cdot)^+$, we can establish the following inequalities:

$$\begin{aligned}
\kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) - \gamma)^+ + \hat{\kappa}(\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \hat{\gamma})^+ &\geq \kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) - \gamma)^+ + \kappa(\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \hat{\gamma})^+ \\
&\geq \kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) + \Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \gamma - \hat{\gamma})^+ \\
&\geq \kappa(\Delta(\boldsymbol{z}+\boldsymbol{\zeta},\hat{\boldsymbol{z}}+\boldsymbol{\zeta}) - \gamma - \hat{\gamma})^+ \\
&= \kappa(\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) - \gamma - \hat{\gamma})^+
\end{aligned}$$

where the last inequality is due to the subadditive property of the coherent disparity metric and the final equality is due to the translation invariant property. Hence, we have

$$\Phi(\boldsymbol{z}) = \inf_{\boldsymbol{\zeta}\in\hat{\mathcal{Z}}} \left\{\kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta}) - \gamma)^+ + \hat{\kappa}(\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) - \hat{\gamma})^+\right\} \geq \kappa(\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) - \gamma - \hat{\gamma})^+.$$

To show that equality holds, we examine the solution of $\boldsymbol{\zeta} \in \hat{\mathcal{Z}}$ under two mutually exclusive and collectively exhaustive cases wherein the minimum can be attained. For the first case of $\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) > \gamma + \hat{\gamma}$, we consider the solution:

$$\boldsymbol{\zeta} = \hat{\boldsymbol{z}} + \frac{\hat{\gamma}}{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})}(\boldsymbol{z} - \hat{\boldsymbol{z}}) = \frac{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) - \hat{\gamma}}{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})}\hat{\boldsymbol{z}} + \frac{\hat{\gamma}}{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})}\boldsymbol{z} \in \mathcal{Z} \subseteq \hat{\mathcal{Z}}.$$

Observe that

$$\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) = \Delta(\boldsymbol{\zeta} - \hat{\boldsymbol{z}},\boldsymbol{0}) = \frac{\hat{\gamma}}{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})}\Delta(\boldsymbol{z} - \hat{\boldsymbol{z}},\boldsymbol{0}) = \hat{\gamma},$$

and because $\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) = \Delta(\boldsymbol{z} - \hat{\boldsymbol{z}},\boldsymbol{0}) > \hat{\gamma}$, we have

$$\Delta(\boldsymbol{z},\boldsymbol{\zeta}) = \Delta(\boldsymbol{z} - \boldsymbol{\zeta},\boldsymbol{0}) = \Delta\left(\left(1 - \frac{\hat{\gamma}}{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})}\right)(\boldsymbol{z} - \hat{\boldsymbol{z}}),\boldsymbol{0}\right) = \Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) - \hat{\gamma}.$$

Hence,

$$\kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta})-\gamma)^+ + \hat{\kappa}(\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}})-\hat{\gamma})^+ = \kappa(\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})-\gamma-\hat{\gamma}).$$

Likewise, for the second case of $\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) \leq \gamma+\hat{\gamma}$, we consider the solution:

$$\boldsymbol{\zeta} = \hat{\boldsymbol{z}} + \frac{\hat{\gamma}}{\gamma+\hat{\gamma}}(\boldsymbol{z}-\hat{\boldsymbol{z}}) = \frac{\hat{\gamma}}{\gamma+\hat{\gamma}}\boldsymbol{z} + \frac{\gamma}{\gamma+\hat{\gamma}}\hat{\boldsymbol{z}} \in \mathcal{Z} \subseteq \hat{\mathcal{Z}}.$$

Observe that

$$\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}}) = \frac{\hat{\gamma}}{\gamma+\hat{\gamma}}\Delta(\boldsymbol{z},\hat{\boldsymbol{z}}) \leq \hat{\gamma} \ \text{ and } \ \Delta(\boldsymbol{z},\boldsymbol{\zeta}) = \Delta(\boldsymbol{z}-\boldsymbol{\zeta},\mathbf{0}) = \Delta\left(\frac{\gamma}{\gamma+\hat{\gamma}}(\boldsymbol{z}-\hat{\boldsymbol{z}}),\mathbf{0}\right) = \gamma\frac{\Delta(\boldsymbol{z},\hat{\boldsymbol{z}})}{\gamma+\hat{\gamma}} \leq \gamma.$$

Hence,

$$\kappa(\Delta(\boldsymbol{z},\boldsymbol{\zeta})-\gamma)^+ + \hat{\kappa}(\Delta(\boldsymbol{\zeta},\hat{\boldsymbol{z}})-\hat{\gamma})^+ = 0.$$

When $\hat{\kappa} < \kappa$, the proof follows a similar analysis and is thus omitted.

□

Proof of Theorem 5.12

Proof. Consider the variable change from $\mathbf{D}^{-1}$ to $\mathbf{Q}$, $\mathbf{D}^{-1}\hat{\boldsymbol{z}}_0$ to $\boldsymbol{q}$, and $\mathbf{D}^{-1}\hat{\mathbf{Z}}$ to $\mathbf{P}$. Given that ψ is a convex and non-increasing function, $\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}}-\boldsymbol{q}-\mathbf{P}\tilde{\boldsymbol{s}}\|)\right]$ is a convex function with respect to the decision variable $(\boldsymbol{q},\mathbf{P},\mathbf{Q})$. The remaining constraints are reformulated in the same way as shown in Theorem 5.8.

□

Proof of Theorem 5.13

Proof. Changing decision variable from $\mathbf{D}^{-2}$ to $\mathbf{W}$, Problem (5.16) is equivalent to

$$\begin{array}{ll} \sup & \log\det(\mathbf{W}) \\ \text{s.t.} & \displaystyle\min_{\hat{\boldsymbol{z}}_0\in\mathbb{R}^L,\,\hat{\mathbf{Z}}\in\mathbb{R}^{L\times S}} \mathbb{E}_{\hat{\mathbb{P}}}\left[\|\mathbf{W}^{1/2}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}_0-\hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\|_2^2\right] \leq L \\ & \mathbf{W} \in \mathbb{S}^L_{++}. \end{array} \tag{C.25}$$

Note that $\mathbb{E}_{\hat{\mathbb{P}}}\left[\|\mathbf{W}^{1/2}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}_0-\hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\|_2^2\right] = \mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}_0-\hat{\mathbf{Z}}\tilde{\boldsymbol{s}})^\top\mathbf{W}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}_0-\hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\right]$ is a quadratic function of $\hat{\boldsymbol{z}}_0$ and $\hat{\mathbf{Z}}$. Since $\mathbf{W} \succ \mathbf{0}$, equating the gradients with respect to $\hat{\boldsymbol{z}}_0$ and $\hat{\mathbf{Z}}$ then yields

$$2\mathbb{E}_{\hat{\mathbb{P}}}\left[\mathbf{W}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{z}}_0-\hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\right] = \mathbf{0} \quad\Longrightarrow\quad \hat{\boldsymbol{z}}_0^\star = \hat{\boldsymbol{\mu}}_{\boldsymbol{z}} - \hat{\mathbf{Z}}^\star\hat{\boldsymbol{\mu}}_{\boldsymbol{s}},$$

and

$$-2\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}_0^\star - \hat{\mathbf{Z}}^\star \tilde{\boldsymbol{s}})\tilde{\boldsymbol{s}}^\top\right] = 0 \quad \Longrightarrow \quad \hat{\boldsymbol{\Sigma}}_{\boldsymbol{z},\boldsymbol{s}} = \hat{\mathbf{Z}}^\star \hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}.$$

Hence, $\hat{\mathbf{Z}}^\star = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}$ and $\hat{\boldsymbol{z}}_0^\star = \hat{\boldsymbol{\mu}}_{\boldsymbol{z}} - \hat{\mathbf{Z}}^\star \hat{\boldsymbol{\mu}}_{\boldsymbol{s}}$. Consequently,

$$\begin{aligned}
\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}_0^\star - \hat{\mathbf{Z}}^\star\tilde{\boldsymbol{s}})^\top \mathbf{W}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}_0^\star - \hat{\mathbf{Z}}^\star\tilde{\boldsymbol{s}})\right] &= \mathbb{E}_{\hat{\mathbb{P}}}\left[\mathrm{tr}(\mathbf{W}(\tilde{\boldsymbol{z}} - \boldsymbol{\mu}_{\boldsymbol{z}} - \hat{\mathbf{Z}}^\star(\tilde{\boldsymbol{s}} - \boldsymbol{\mu}_{\boldsymbol{s}}))(\tilde{\boldsymbol{z}} - \boldsymbol{\mu}_{\boldsymbol{z}} - \hat{\mathbf{Z}}^\star(\tilde{\boldsymbol{s}} - \boldsymbol{\mu}_{\boldsymbol{s}}))^\top)\right] \\
&= \mathrm{tr}(\mathbf{W}(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\mathbf{Z}}^\star\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}(\hat{\mathbf{Z}}^\star)^\top + \hat{\mathbf{Z}}^\star\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}(\hat{\mathbf{Z}}^\star)^\top)) \\
&= \mathrm{tr}(\mathbf{W}(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})).
\end{aligned}$$

Problem (C.25) then becomes to

$$\begin{array}{ll}
\sup & \log\det(\mathbf{W}) \\
\text{s.t.} & \mathrm{tr}(\mathbf{W}(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})) \leq L \\
& \mathbf{W} \in \mathbb{S}_{++}^L.
\end{array} \tag{C.26}$$

The Lagrangian function for this Problem is:

$$\Lambda(\mathbf{W}, \lambda) = \log\det(\mathbf{W}) - \lambda\,\mathrm{tr}(\mathbf{W}(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})) + \lambda L.$$

We can calculate the gradient with respect $\mathbf{W}$ as follows:

$$\nabla_{\mathbf{W}}\Lambda(\mathbf{W}, \lambda) = \mathbf{W}^{-1} - \lambda(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}}).$$

Equating it to zero, the optimal solution shall satisfy:

$$\lambda^\star\mathbf{W}^\star = (\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})^{-1}.$$

Complementary slackness of the KKT conditions requires that

$$\lambda^\star(\mathrm{tr}(\mathbf{W}(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})) - L) = 0,$$

implying $\lambda^\star = 1$. Therefore, we obtain the optimal solution of the deviation matrix via

$$\mathbf{W}^\star = (\hat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{sz}})^{-1} \text{ and } \mathbf{D}^\star = (\mathbf{W}^\star)^{-\frac{1}{2}}.$$

The optimality is ensured, as Problem (C.26) is a convex optimization problem.

$\square$

Proof of Theorem 5.14

Proof. The square loss function can be reformulated as follows:

$$\begin{aligned}\mathbb{E}_{\hat{\mathbb{P}}}\left[\|\hat{\mathbf{D}}^{-1}(\boldsymbol{\zeta}_0+\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})\|_2^2\right] &= \mathbb{E}_{\hat{\mathbb{P}}}\left[(\boldsymbol{\zeta}_0+\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}(\boldsymbol{\zeta}_0+\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})\right]\\ &= \mathbb{E}_{\hat{\mathbb{P}}}\left[\boldsymbol{\zeta}_0^\top\hat{\mathbf{D}}^{-2}\boldsymbol{\zeta}_0+2\boldsymbol{\zeta}_0^\top\hat{\mathbf{D}}^{-2}(\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})+(\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}(\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})\right]\\ &= \boldsymbol{\zeta}_0^\top\hat{\mathbf{D}}^{-2}\boldsymbol{\zeta}_0+2\boldsymbol{\zeta}_0^\top\hat{\mathbf{D}}^{-2}\mathbb{E}_{\hat{\mathbb{P}}}\left[\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}}\right]+\mathbb{E}_{\hat{\mathbb{P}}}\left[(\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}(\mathbf{Z}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})\right],\end{aligned}$$

which, for any fixed $\mathbf{Z}$, is a quadratic function of $\boldsymbol{\zeta}_0$. Minimizing this quadratic function with respect to $\boldsymbol{\zeta}_0$ leads to the following optimal map:

$$\boldsymbol{\zeta}_0^\star(\mathbf{Z}) = -\mathbf{Z}\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{s}}\right]+\mathbb{E}_{\hat{\mathbb{P}}}\left[\tilde{\boldsymbol{z}}\right] = \hat{\boldsymbol{\mu}}_{\boldsymbol{z}}-\mathbf{Z}\hat{\boldsymbol{\mu}}_{\boldsymbol{s}};$$

and results in:

$$\begin{aligned}&\mathscr{L}(\boldsymbol{\zeta}_0^\star(\mathbf{Z}),\mathbf{Z};\mathcal{D})\\ =\;& \mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{s}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})^\top\mathbf{Z}^\top\hat{\mathbf{D}}^{-2}\mathbf{Z}(\tilde{\boldsymbol{s}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})-2(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}\mathbf{Z}(\tilde{\boldsymbol{s}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})\right]\\ &+\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{z}})\right]\\ =\;& \operatorname{tr}(\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{s}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})(\tilde{\boldsymbol{s}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})^\top\right]\mathbf{Z}^\top\hat{\mathbf{D}}^{-2}\mathbf{Z})-2\operatorname{tr}(\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{s}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}})^\top\right]\hat{\mathbf{D}}^{-2}\mathbf{Z})\\ &+\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}})^\top\hat{\mathbf{D}}^{-2}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{z}})\right]\\ =\;& \operatorname{tr}(\mathbf{Z}^\top\hat{\mathbf{D}}^{-2}\mathbf{Z}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}})-2\operatorname{tr}(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}^\top\hat{\mathbf{D}}^{-2}\mathbf{Z})+\mathbb{E}_{\hat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}}_{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}(\tilde{\boldsymbol{z}}-\hat{\boldsymbol{\mu}})\right].\end{aligned}\tag{C.27}$$

The gradient with respect to $\mathbf{Z}$ is

$$\nabla_{\mathbf{Z}}\mathscr{L}(\boldsymbol{\zeta}_0^\star(\mathbf{Z}),\mathbf{Z};\mathcal{D}) = 2\hat{\mathbf{D}}^{-2}\mathbf{Z}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}-2\hat{\mathbf{D}}^{-2}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}.$$

As $\hat{\mathbf{D}}$ and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}$ are invertible, equating the gradient to zero yields a solution

$$\mathbf{Z}^\star = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}}^{-1},$$

which is optimal as Equation (C.27), with the first term being quadratic and the second term linear, is convex in $\mathbf{Z}$. As a result, the OLS estimators are $\hat{\mathbf{Z}} = \mathbf{Z}^\star$ and $\hat{\boldsymbol{z}}_0 = \boldsymbol{\zeta}_0^\star(\mathbf{Z}^\star)$.

Consequently, we have

$$\mathscr{L}(\hat{\boldsymbol{z}}_0,\hat{\mathbf{Z}};\mathcal{D}) = -\hat{\boldsymbol{z}}_0^\top\hat{\mathbf{D}}^{-2}\hat{\boldsymbol{z}}_0+\mathbb{E}_{\hat{\mathbb{P}}}\left[(\hat{\mathbf{Z}}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})^\top\hat{\mathbf{D}}^{-2}(\hat{\mathbf{Z}}\tilde{\boldsymbol{s}}-\tilde{\boldsymbol{z}})\right].$$

Hence, the estimation disparity metric associated with the square loss function simplifies to:

$$\begin{aligned}
\hat{\Delta}((\boldsymbol{u}, \mathbf{U}), (\mathbf{0}, \mathbf{0})) &= \eta(\mathscr{L}(\hat{\boldsymbol{z}}_0 + \boldsymbol{u}, \hat{\mathbf{Z}} + \mathbf{U}; \mathcal{D}) - \mathscr{L}(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}}; \mathcal{D})) \\
&= \eta(\mathbb{E}_{\hat{\mathbb{P}}}\,[\|\hat{\mathbf{D}}^{-1}(\hat{\boldsymbol{z}}_0 + \boldsymbol{u} + (\hat{\mathbf{Z}} + \mathbf{U})\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{z}})\|_2^2] - \mathscr{L}(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}}; \mathcal{D})) \\
&= \eta\Big(\mathbb{E}_{\hat{\mathbb{P}}}\,[\|\hat{\mathbf{D}}^{-1}(\boldsymbol{u} + \mathbf{U}\tilde{\boldsymbol{s}})\|_2^2] + 2\mathbb{E}_{\hat{\mathbb{P}}}\,[(\boldsymbol{u} + \mathbf{U}\tilde{\boldsymbol{s}})^\top \hat{\mathbf{D}}^{-2}(\hat{\boldsymbol{z}}_0 + \hat{\mathbf{Z}}\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{z}})] \\
&\qquad + \mathbb{E}_{\hat{\mathbb{P}}}\,[\|\hat{\mathbf{D}}^{-1}(\hat{\boldsymbol{z}}_0 + \hat{\mathbf{Z}}\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{z}})\|_2^2] - \mathscr{L}(\hat{\boldsymbol{z}}_0, \hat{\mathbf{Z}}; \mathcal{D})\Big) \\
&= \eta(\|\hat{\mathbf{D}}^{-1}(\boldsymbol{u} + \mathbf{U}\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})\|_2^2 + \mathbb{E}_{\hat{\mathbb{P}}}\,[(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}})^\top \mathbf{U}^\top \hat{\mathbf{D}}^{-2}\mathbf{U}(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}})]) \\
&= \eta(\|\hat{\mathbf{D}}^{-1}(\boldsymbol{u} + \mathbf{U}\hat{\boldsymbol{\mu}}_{\boldsymbol{s}})\|_2^2 + \mathrm{tr}(\mathbf{U}^\top \hat{\mathbf{D}}^{-2}\mathbf{U}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}})),
\end{aligned}$$

where in the fourth equation, we have that

$$\begin{aligned}
\mathbb{E}_{\hat{\mathbb{P}}}\,[(\boldsymbol{u} + \mathbf{U}\tilde{\boldsymbol{s}})^\top \hat{\mathbf{D}}^{-2}(\hat{\boldsymbol{z}}_0 + \hat{\mathbf{Z}}\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{z}})] &= \mathbb{E}_{\hat{\mathbb{P}}}\,[(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}})^\top \mathbf{U}^\top \hat{\mathbf{D}}^{-2}(\hat{\mathbf{Z}}(\tilde{\boldsymbol{s}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{s}}) - (\tilde{\boldsymbol{z}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{z}}))] \\
&= \mathrm{tr}(\mathbf{U}^\top \hat{\mathbf{D}}^{-2}(\hat{\mathbf{Z}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{s}} - \hat{\boldsymbol{\Sigma}}_{\boldsymbol{zs}})) \\
&= 0.
\end{aligned}$$

The proof is then completed.

□

C.2 Variance of input vector estimation

We consider a symmetric outcome confidence set without incorporating side information. Consider a random input vector $\tilde{\boldsymbol{z}}$ expressed as $\tilde{\boldsymbol{z}} = \boldsymbol{\Sigma}^{\frac{1}{2}}\tilde{\boldsymbol{v}}$ that involves some i.i.d. random factors $\tilde{v}_\ell$, $\ell \in [L]$, with $L = 10$. Here, $\tilde{v}_\ell$ follows either a standard normal distribution $\mathbb{N}(0, 1)$ or a uniform distribution $\mathbb{U}[0, \sqrt{12}]$ of the same mean 0 and variance 1. The positive definite matrix $\boldsymbol{\Sigma}$ introduces correlations among the random components $\tilde{z}_\ell$, $\ell \in [L]$ but ensures that the variance of each component is 100.

$$\boldsymbol{\Sigma} = \begin{pmatrix}
100.00 & 26.04 & 39.77 & 40.30 & -26.28 & -11.33 & -7.22 & -20.18 & -33.85 & -1.51 \\
26.04 & 100.00 & 51.87 & 26.51 & -40.63 & -9.35 & -30.00 & 5.93 & 8.55 & -3.98 \\
39.77 & 51.87 & 100.00 & 19.47 & -31.73 & -18.13 & -43.93 & 0.84 & -31.34 & -10.16 \\
40.30 & 26.51 & 19.47 & 100.00 & -30.80 & -5.80 & 12.69 & -48.35 & 6.47 & 18.48 \\
-26.28 & -40.63 & -31.73 & -30.80 & 100.00 & 26.96 & 22.66 & 22.25 & -11.01 & -8.11 \\
-11.33 & -9.35 & -18.13 & -5.80 & 26.96 & 100.00 & 51.40 & 40.78 & -25.34 & 11.41 \\
-7.22 & -30.00 & -43.93 & 12.69 & 22.66 & 51.40 & 100.00 & 3.17 & -0.75 & 39.06 \\
-20.18 & 5.93 & 0.84 & -48.35 & 22.25 & 40.78 & 3.17 & 100.00 & -15.75 & 10.53 \\
-33.85 & 8.55 & -31.34 & 6.47 & -11.01 & -25.34 & -0.75 & -15.75 & 100.00 & 27.12 \\
-1.51 & -3.98 & -10.16 & 18.48 & -8.11 & 11.41 & 39.06 & 10.53 & 27.12 & 100.00
\end{pmatrix}.$$

Given a number Ω of historical dataset $\mathcal{D} = \{(1, \boldsymbol{z}_1), \ldots, (\Omega, \boldsymbol{z}_\Omega)\}$, we employ the $\widehat{\text{MVCS}}$ approach to estimate $\hat{\boldsymbol{z}} = (\hat{z}_\ell)_{\ell \in [L]}$. This estimator is a function of historical data and, thereby, is inherently random. To implement the $\widehat{\text{MVCS}}$ problem, we employ the L_{p_1}-norm and a power deviation penalty function $\psi_{p_2}(r) = r^{p_2}$, and the following case uses a linear penalty, *i.e.*, $p_2 = 1$. Additionally, we set the regularization vector $\boldsymbol{r} = \mathbf{1}$ for all values of Ω to avoid the feasibility issue that $\det(\mathbf{D}) = 0$ when $\Omega = L = 10$.

To assess the stability and reliability of the estimates, we consider 200 replications that generate the historical dataset from the underlying $\mathbb{N}(0, 1)$ or $\mathbb{U}[0, \sqrt{12}]$ and compute the following average of variances of $\hat{\boldsymbol{z}}$'s components over these 200 replications:

$$\frac{1}{L} \sum_{\ell \in [L]} \mathbb{V}\text{ar}(\hat{z}_l).$$

We call this average the variance of estimation and report it in Table C.1, wherein the first two columns present results for the uniform distribution and the normal distribution, respectively.

Ω	Uniform w.o. Outliers				Normal w.o. Outliers				Uniform with Outliers				Normal with Outliers			
	L_1	L_2	L_3	L_4	L_1	L_2	L_3	L_4	L_1	L_2	L_3	L_4	L_1	L_2	L_3	L_4
10	23.61	11.33	9.85	9.35	16.30	10.58	11.74	13.00	28.05	48.75	138.39	221.50	26.11	51.09	141.85	226.24
20	13.35	5.81	4.72	4.34	9.18	5.51	5.89	6.62	14.17	17.14	51.66	102.58	10.40	16.23	51.75	100.05
50	5.53	2.13	1.55	1.33	3.07	2.08	2.40	2.83	5.36	4.83	10.43	18.92	3.88	4.82	10.76	18.88
100	2.79	1.11	0.78	0.65	1.64	1.08	1.18	1.35	2.81	2.36	3.55	4.81	1.85	2.22	4.08	5.90

Table C.1: Variance of estimation.

For the uniform distribution, the input vector subjected to the L_3-norm exhibits the lowest variance of estimation. Also, a general trend indicates that the variance of estimation decreases as the parameter p_1 for the L_{p_1}-norm increases. Conversely, for the normal distribution, the input vector subject to the L_2-norm shows the lowest variance. These observations suggest that different probability distributions may favor different norms.

Interestingly, these observations are also consistent with the uniformly minimum-variance unbiased estimator (UMVUE) when the uncertainty is one-dimensional (*i.e.*, $L = 1$). The empirical mean $\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{z}] = \sum_{\omega=1}^{\Omega} z_\omega / \Omega$ is the UMVUE of normal distribution (Rohatgi and Saleh 2015), while the average of the smallest and largest order statistics, denoted as $(\hat{z}_{(1)} + \hat{z}_{(\Omega)})/2$, emerges as the UMVUE of uniform distribution (Bar-lev and Boukai 1985). These estimators corroborate our analytical solutions as delineated in Theorem 5.9, as the empirical mean $\mathbb{E}_{\hat{\mathbb{P}}}[\tilde{z}]$ solves $\mathbb{E}_{\hat{\mathbb{P}}}[(\tilde{z} - \hat{z}^\star)^+] = \mathbb{E}_{\hat{\mathbb{P}}}[(\tilde{z} - \hat{z}^\star)^-]$ while $(\hat{z}_{(1)} + \hat{z}_{(\Omega)})/2$ satisfies $\operatorname{ess\,sup}_{\hat{\mathbb{P}}}(\tilde{z} - \hat{z}^\star) = \operatorname{ess\,sup}_{\hat{\mathbb{P}}}(\hat{z}^\star - \tilde{z})$—the limiting case of $\mathbb{E}_{\hat{\mathbb{P}}}\left[((\tilde{z} - \hat{z}^\star)^+)^{p-1}\right] = \mathbb{E}_{\hat{\mathbb{P}}}\left[((\tilde{z} - \hat{z}^\star)^-)^{p-1}\right]$ as $p \to +\infty$.

We next consider the possibility that the historical dataset may be contaminated by outliers. In particular, we introduce these outliers as drifts 100 (resp., -100) at a probability of 0.025 (resp., 0.025), which are designed to not influence the expectation of the distributions. In other words, the distribution of $\tilde{z}_\ell$ is now a mixture distribution that assigns 0.95-weight to the rotated-by-$\boldsymbol{\Sigma}$ normal or uniform random variables and equally assigns 0.025-weight to the distributions shifted by 100 and -100, respectively. Other procedures remain similar. The results are presented in the third and fourth columns of Table C.1. It can be seen that the $\widehat{\text{MVCS}}$ approach is more sensitive to outliers as the parameter p_1 for the L_{p_1}-norm decreases.

C.3 Additional details

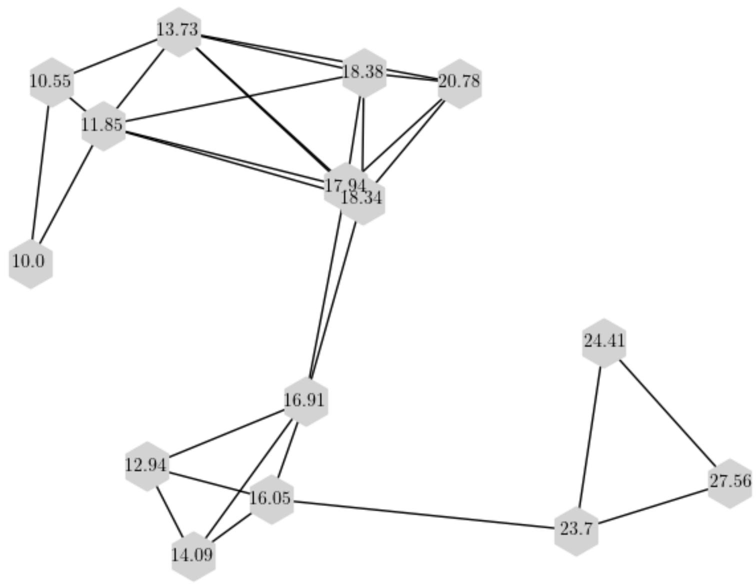


Figure C.8: Resource allocation network: 15 nodes and 30 arcs.

Notes. The parameter μ_i for each normally distributed demand is indicated at the node.

We present the replication-by-replication comparison of RS (resp., RO) with EO in Figure C.10 (resp., Figure C.9), wherein each point corresponds to a specific replication. The RS model, employing both L_1- and L_2-norms, demonstrates superior performance compared to EO in the majority of replications. Additionally, the RO model, across all norms, shows superior performance compared to EO in the majority of replications.

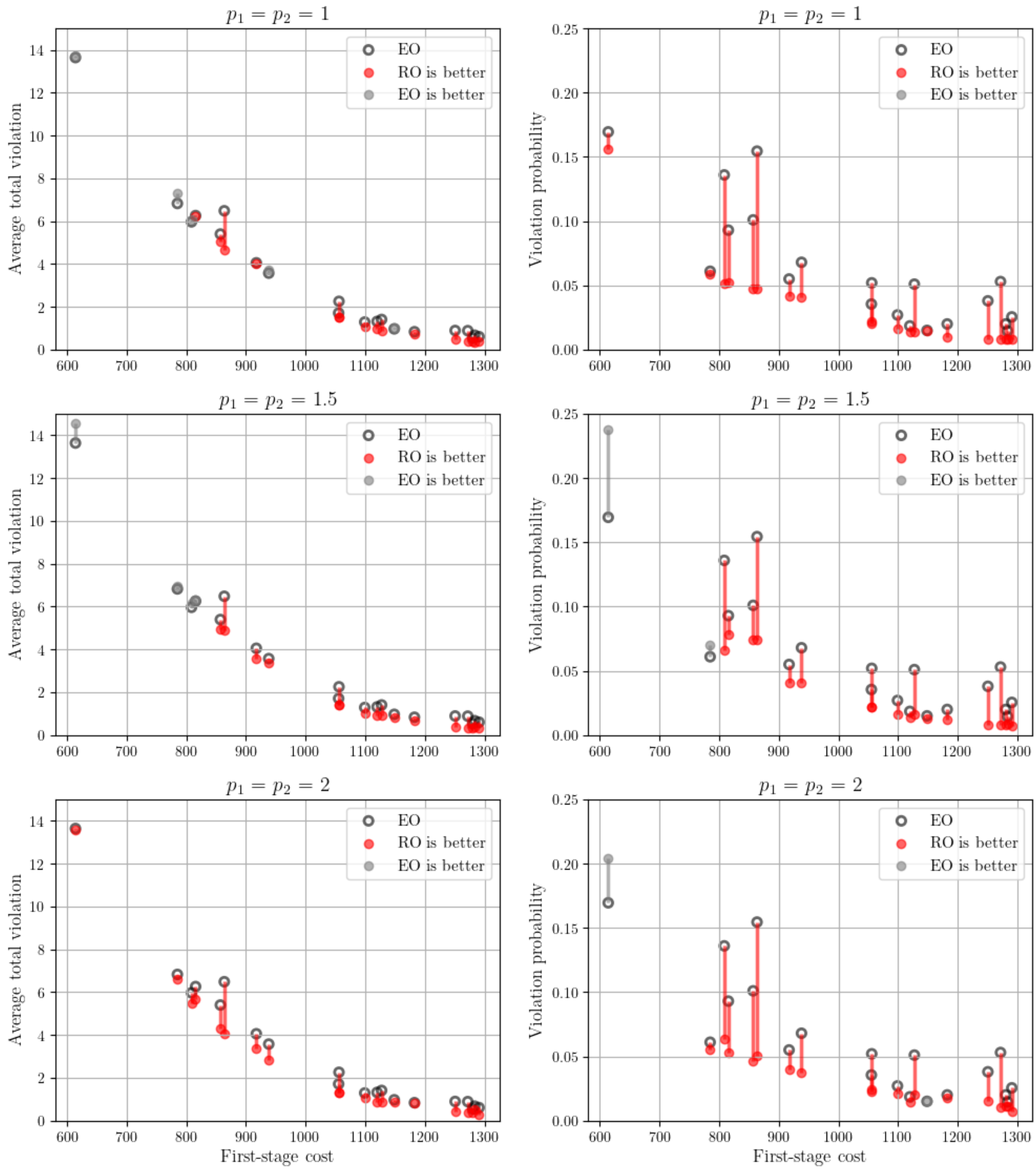


Figure C.9: Infeasibility profile of EO and RO models with samples $\Omega = 40$ in each replication.

Notes. The grey color indicates that EO performs better in view of the specific performance metric, while the red color indicates that RO performs better.

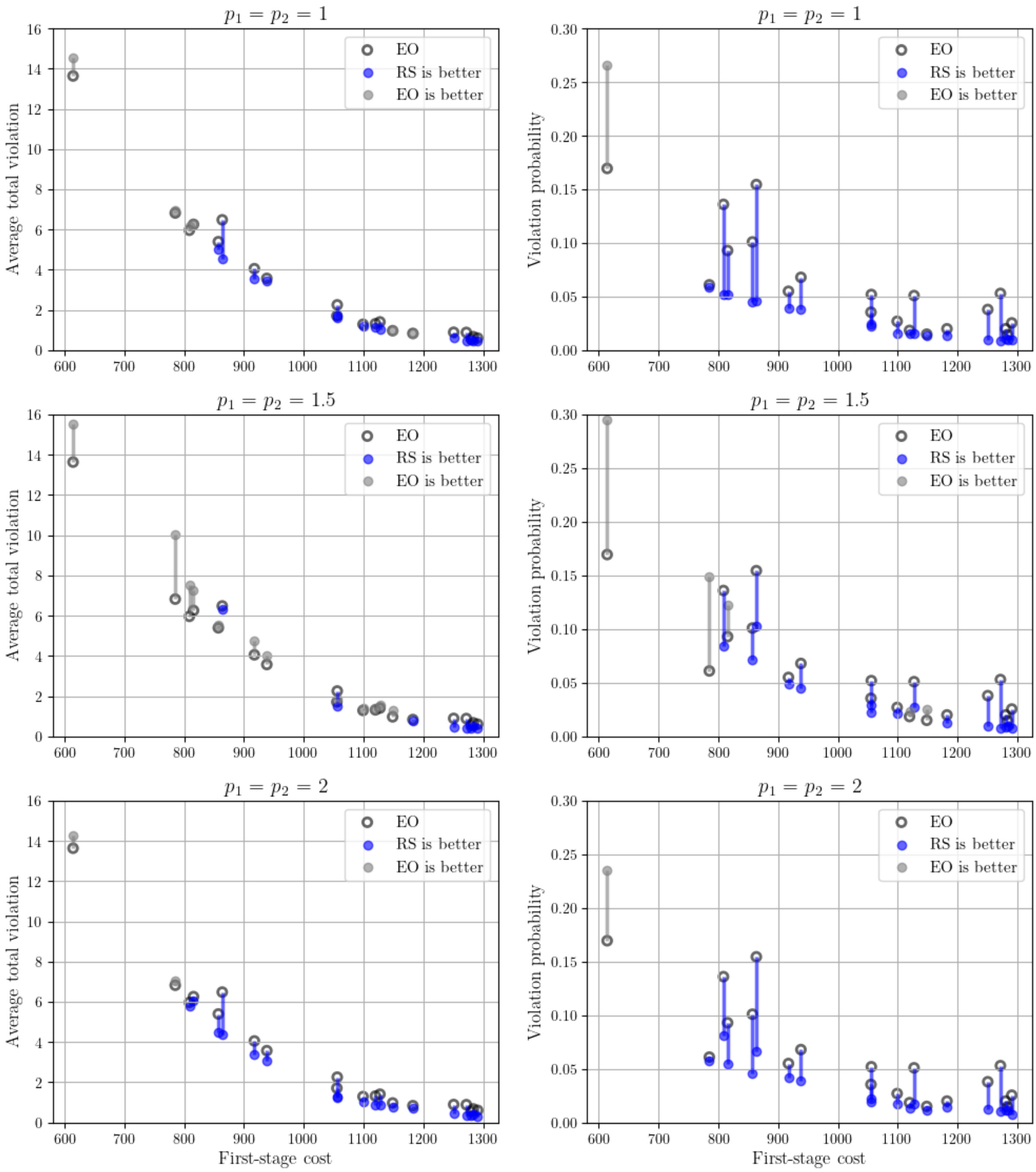


Figure C.10: Infeasibility profile of EO and RS models with samples $\Omega = 40$ in each replication.

Notes. The grey color indicates that EO performs better in view of the specific performance metric, while the blue color indicates that RS performs better.

C.4 Sample code of numerical cases

The emergency service resource allocation problem described in Section 5.6 is solved using the RSOME package. The sample code for running the robust satisficing model is provided below as a guideline to implement the proposed methods.

```
import numpy as np
import rsome as rso
from rsome import ro
from rsome import cpt_solver as cpt

def mvcs(z, p1=2, p2=2, display=True):
    """
    Return the vector q and the symmetrical matrix Q as the results of estimating
    the minimum volume confidence set (MVCS), following Eq. (9).

    Parameters
    ----------
    z : np.ndarray
        The training dataset is given as a 2-D array.
    p1 : int
        Specify the Lp-norm of the MVCS estimation. The parameter p1 must be
        an integer larger than one. The default value is two.
    p2 : int
        Specify the p-value of the power deviation penalty function. The parameter
        p2 must be an integer larger than one. See Definition 4.2. The default
        value is two.
    display: bool
        If the log information is displayed.

    Return
    ------
    q : np.ndarray
        A 1-D array for the constant term of the MVCS approach.
    Q : np.ndarray
        A 2-D array for the linear term of the MVCS approach.
    """

    N, L = z.shape

    model = ro.Model()
    q = model.dvar(L)
    u = model.dvar(N)
    v = model.dvar(N)
```

```
    Q = model.dvar((L, L))

    model.max(rso.rootdet(Q))
    model.st((1/N) * u.sum() <= 1)
    model.st((1/L) * rso.power(v, p2) <= u)
    l_norm = lambda x: rso.norm(x, p1)
    for n in range(N):
        model.st(l_norm(Q@z[n] - q) <= v[n])

    if display:
        msg = f"Sample data:          {N} records x {L} inputs\n"
        msg += f"Norm type:            l{p1}-norm\n"
        msg += f"Deviation penalty: power={p2}\n"
        print(msg)
        print(model.do_math())

    model.solve(cpt, display=display)

    return q.get(), Q.get()


def rs_decision(c, arcs, xbar, tau, zin, p1=2, p2=2, display=True):
    """
    Return the objective value and optimal solution of the robust satisficing
    model for the emergency service resource allocation problem.

    Parameters
    ----------
    c : np.ndarray
        A 1-D array representing the per-unit costs of emergency service
        resources.
    arcs: list
        A list of pairs of node indices, representing the routes between
        locations where the service can be transshipped.
    xbar : float
        The maximum resource capacity of each node.
    tau : float
        The budget of first-stage cost, in terms of the total cost of
        inventory at all locations.
    zin : np.ndarray
        The training dataset is given as a 2-D array.
    p1 : int
        Specify the Lp-norm of the MVCS estimation. The parameter p1 must be
        an integer larger than one. The default value is two.
```

```
    p2 : int
        Specify the p-value of the power deviation penalty function.
        The parameter p2 must be an integer larger than one.
        See Definition 4.2. The default value is two.
    display: bool
        If the log information is displayed.

    Return
    ------
    objval : float
        The optimal objective value of the robust satisficing model
    xs : np.ndarray
        A 1-D array representing the optimal solution of the robust
        satisficing model.
    """

    I = len(c)
    ij= np.array([(i, i) for i in range(I)] +
                 arcs + [a[::-1] for a in arcs])
    K = ij.shape[0]

    q, Q = mvcs(zin, p1, p2, display)
    gbar = np.linalg.norm(zin@Q - q, p1, axis=1).max()

    model = ro.Model()
    z = model.rvar(I)
    w1 = model.rvar(I)
    w2 = model.rvar(I)
    u = model.rvar()
    zset = (w1 - w2 == Q@z - q,
            rso.norm(w1 + w2, p1) <= u, u <= gbar,
            w1 >= 0, w2 >= 0,
            z >= 0, z <= 100)

    x = model.dvar(I)
    kappa = model.dvar()
    y = model.ldr(K)
    y.adapt(w1)
    y.adapt(w2)
    y.adapt(u)

    model.minmax(kappa, zset)
    model.st(c @ x <= tau)
    for i in range(I):
```

```
127             model.st(y[ij[:, 0] == i].sum() <= x[i])
128             model.st(z[i] - y[ij[:, 1] == i].sum() <= kappa * u)
129         model.st(x >= 0, x <= xbar, y >= 0)
130         model.solve(cpt, display=display)
131 
132         objval, xs = model.get(), x.get()
133         return objval, xs
```

In the code above, the function `rs_decision` is utilized to formulate and solve the robust satisficing model, where the function `mvcs` defined at the top is called in line 104 to solve the $\widehat{\text{MVCS}}$ problem and return parameters of the uncertainty set.

6 Conclusion

This dissertation presents three complementary contributions to the field of robust decision-making under uncertainty, with a particular focus on applications in healthcare operations and public health. The developed models address challenges at different hierarchical levels: patient-level treatment decisions, population-level pandemic interventions, and resource allocation in emergency medical services. Each contribution is motivated by real-world problems characterized by both outcome uncertainty and estimation uncertainty, two factors that critically affect the reliability and effectiveness of data-driven decision-making in healthcare.

In Chapter 3, we develop a risk-sensitive optimal stopping framework for sequential medical decision-making in intensive care settings. By integrating the entropic risk measure into a MDP model, the framework accounts not only for the expected length of stay but also for variability and tail risks in clinical outcomes. Structural analysis reveals that the optimal policies possess a threshold structure, providing interpretability and operational relevance. Additionally, we integrate the predictive information into our ROSP model and analyze the corresponding structure properties. The numerical results, based on extubation and discharge decisions, demonstrate that the risk-sensitive policies yield a more robust trade-off between patient safety and resource utilization compared to traditional risk-neutral models.

In Chapter 4, attention is shifted to the public health domain, where we address the problem of pandemic intervention planning under epidemiological uncertainty. We proposed a robust decision-making model based on stochastic compartmental dynamics, introducing an ambiguity tolerance metric to quantify the violation of healthcare performance targets. The framework explicitly models the impact of interventions on disease evolution and healthcare capacity. To enhance tractability, we utilize the asymptotic Gaussian property of compartment populations and derive a bilinear approximation of the robust model. The application to real-world data from Singapore during the COVID-19 pandemic confirms that the proposed approach outperforms conventional deterministic models, offering a more balanced and resilient intervention strategy.

In Chapter 5, we proposed a unified robust decision-making framework for resource allocation problems, particularly in the context of EMS. The framework systematically incorporates two layers of uncertainty: *outcome uncertainty*, describing the deviation between realized and specified inputs, and *estimation uncertainty*, accounting for inaccuracies in input parameter estimation. We introduced the MVCS method to construct norm-specific uncertainty sets from historical data. The approach is extended to handle asymmetries and to incorporate side information through predictive models. Numerical experiments show that the unified framework yields less conservative but more reliable solutions, especially in data-scarce settings and in the presence of outliers.

Taken together, the results of this dissertation contribute to the methodological advancement of robust optimization by integrating robust satisficing principles and risk-sensitive modeling into a coherent framework. The presented approaches offer practical tools for improving decision-making robustness in healthcare and public health contexts. They allow for better handling of uncertainty, more interpretable policies, and enhanced performance even under imperfect data conditions.

Future research may further extend these models by considering additional forms of uncertainty, adaptive decision-making under evolving information, and the integration of human-in-the-loop systems for healthcare and public sector applications.

7 Bibliography

Ahipaşaoğlu, S. D. (2015). Fast algorithms for the minimum volume estimator. In: *Journal of Global Optimization* **62**, pp. 351–370.

Ajdari, A., Niyazi, M., Nicolay, N. H., Thieke, C., Jeraj, R. and Bortfeld, T. (2019). Towards optimal stopping in radiation therapy. In: *Radiotherapy and Oncology* **134**, pp. 96–100.

Alagoz, O., Maillart, L. M., Schaefer, A. J. and Roberts, M. S. (2004). The optimal timing of living-donor liver transplantation. In: *Management Science* **50**(10), pp. 1420–1430.

Atamtürk, A. and Zhang, M. (2007). Two-stage robust network flow and design under demand uncertainty. In: *Operations Research* **55**(4), pp. 662–673.

Aumann, R. J. and Serrano, R. (2008). An economic index of riskiness. In: *Journal of Political Economy* **116**(5), pp. 810–836.

Ayvaci, M. U., Alagoz, O. and Burnside, E. S. (2012). The Effect of Budgetary Restrictions on Breast Cancer Diagnostic Decisions. In: *Manufacturing & Service Operations Management* **14**(4), p. 600.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**(3), pp. 579–602.

Bandi, C. and Bertsimas, D. (2012). Tractable stochastic analysis in high dimensions via robust optimization. In: *Mathematical Programming* **134**, pp. 23–70.

Bar-lev, S. and Boukai, B. (1985). Minimum variance unbiased estimation for families of distributions involving two truncation parameters. In: *Journal of Statistical Planning and Inference* **12**, pp. 379–384.

Barrett, M., Smith, M., Elixhauser, A., Honigman, L. and Pines, J. (2014). Utilization of intensive care services, 2011. In: *HCUP statistical brief* **185**.

Bäuerle, N. and Ott, J. (2011). Markov decision processes with average-value-at-risk criteria. In: *Mathematical Methods of Operations Research* **74**, pp. 361–379.

Bäuerle, N. and Rieder, U. (2011). *Markov decision processes with applications to finance.* Springer Science & Business Media.

— (2014). More risk-sensitive Markov decision processes. In: *Mathematics of Operations Research* **39**(1), pp. 105–120.

Bellman, R. (1957). A Markovian decision process. In: *Journal of mathematics and mechanics*, pp. 679–684.

Ben-Tal, A., Bertsimas, D. and Brown, D. (2010). A soft robust model for optimization under ambiguity. In: *Operations Research* **58**, pp. 1220–1234.

Ben-Tal, A., Brekelmans, R., Hertog, D. den and Vial, J. (2017). Globalized robust optimization for nonlinear uncertain inequalities. In: *INFORMS Journal on Computing* **29**(2), pp. 350–366.

Ben-Tal, A., Goryashko, A., Guslitzer, E. and Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. In: *Mathematical programming* **99**(2), pp. 351–376.

Ben-Tal, A., Hertog, D. den, De Waegenaere, A., Melenberg, B. and Rennen, G. (2013). Robust solutions of optimization problems affected by uncertain probabilities. In: *Management Science* **59**(2), pp. 341–357.

Ben-Tal, A. and Nemirovski, A. (1998). Robust convex optimization. In: *Mathematics of operations research* **23**(4), pp. 769–805.

Bennouna, A., Joseph, J., Nze-Ndong, D., Perakis, G., Singhvi, D., Lami, O. S., Spantidakis, Y., Thayaparan, L. and Tsiourvas, A. (2023). COVID-19: Prediction, prevalence, and the operations of vaccine allocation. In: *Manufacturing & Service Operations Management* **25**(3), pp. 1013–1032.

Beraldi, P., Bruni, M. E. and Conforti, D. (2004). Designing robust emergency medical service via stochastic programming. In: *European Journal of Operational Research* **158**(1), pp. 183–193.

Bertsimas, D., Boussioux, L., Cory-Wright, R., Delarue, A., Digalakis, V., Jacquillat, A., Kitane, D. L., Lukin, G., Li, M., Mingardi, L., et al. (2021a). From predictions to prescriptions: A data-driven response to COVID-19. In: *Health care management science*, pp. 1–20.

Bertsimas, D., den Hertog, D. and Pauphilet, J. (2021b). Probabilistic guarantees in robust optimization. In: *SIAM Journal on Optimization* **31**(4), pp. 2893–2920.

Bertsimas, D., Digalakis Jr, V., Jacquillat, A., Li, M. L. and Previero, A. (2022). Where to locate COVID-19 mass vaccination facilities? In: *Naval Research Logistics* **69**(2), pp. 179–200.

Bertsimas, D., Gupta, V. and Kallus, N. (2018). Data-driven robust optimization. In: *Mathematical Programming* **167**(2), pp. 235–292.

Bertsimas, D., McCord, C. and Sturt, B. (2023). Dynamic optimization with side information. In: *European Journal of Operational Research* **304**(2), pp. 634–651.

Bertsimas, D., Pachamanova, D. and Sim, M. (2004). Robust linear optimization under general norms. In: *Operations Research Letters* **32**(6), pp. 510–516.

Bertsimas, D. and Sim, M. (2004). The price of robustness. In: *Operations research* **52**(1), pp. 35–53.

— (2006). Tractable approximations to robust conic optimization problems. In: *Mathematical programming* **107**(1), pp. 5–36.

Bertsimas, D. and Van Parys, B. (2022). Bootstrap robust prescriptive analytics. In: *Mathematical Programming* **195**(1), pp. 39–78.

Boutilier, J. and Chan, T. (2020). Ambulance emergency response optimization in developing countries. In: *Operations Research* **68**(5), pp. 1315–1334.

Boyarchenko, S. and Levendorskii, S. (2007). *Irreversible decisions under uncertainty: optimal stopping made easy*. Vol. 27. Springer Science & Business Media.

Boyd, S. P. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Brown, D. B. and Sim, M. (2009). Satisficing measures for analysis of risky positions. In: *Management Science* **55**(1), pp. 71–84.

Calafiore, G. and Campi, M. (2005). Uncertain convex programs: Randomized solutions and confidence levels. In: *Mathematical Programming* **102**, pp. 25–46.

Chan, C. W., Farias, V. F., Bambos, N. and Escobar, G. J. (2012). Optimizing intensive care unit discharge decisions with patient readmissions. In: *Operations Research* **60**(6), pp. 1323–1341.

Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D. and Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. In: *Nature* **589**(7840), pp. 82–87.

Chen, X., Sim, M. and Sun, P. (2007). A robust optimization perspective on stochastic programming. In: *Operations research* **55**(6), pp. 1058–1071.

Chen, Z., Chua, J., Sim, M. and Xiong, P. (2023). The Occam's Razor of Robustness. In: *SSRN 4590551*.

Chen, Z., Sim, M. and Xiong, P. (2020). Robust stochastic optimization made easy with RSOME. In: *Management Science* **66**(8), pp. 3329–3339.

Chen, Z. and Xiong, P. (2023). RSOME in Python: An open-source package for robust stochastic optimization made easy. In: *INFORMS Journal on Computing* **35**(4), pp. 717–724.

Cheng, G., Xie, J., Zheng, Z., Luo, H. and Ooi, O. C. (2024). Extubation decision making with predictive information for mechanically ventilated patients in ICU. In: *Management Science*. **Forthcoming**.

Chhatwal, J., Alagoz, O. and Burnside, E. S. (2010). Optimal breast biopsy decision-making based on mammographic features and demographic factors. In: *Operations Research* **58**(6), pp. 1577–1591.

Chung, K. L. (2001). *A course in probability theory*. Academic press.

Ciocan, D. F. and Mišić, V. V. (2022). Interpretable optimal stopping. In: *Management Science* **68**(3), pp. 1616–1638.

Cohen, I. G., Crespo, A. M. and White, D. B. (2020). Potential legal liability for withdrawing or withholding ventilators during COVID-19: assessing the risks and identifying needed reforms. In: *Jama* **323**(19), pp. 1901–1902.

DeMiguel, V., Garlappi, L. and Uppal, R. (2009). Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? In: *The Review of Financial Studies* **22**(5), pp. 1915–1953.

Denton, B. T., Kurt, M., Shah, N. D., Bryant, S. C. and Smith, S. A. (2009). Optimizing the start time of statin therapy for patients with diabetes. In: *Medical Decision Making* **29**(3), pp. 351–367.

Donsker, M. D. and Varadhan, S. S. (1983). Asymptotic evaluation of certain Markov process expectations for large time. IV. In: *Communications on pure and applied mathematics* **36**(2), pp. 183–212.

Duijzer, L., Van Jaarsveld, W. and Dekker, R. (2018a). Literature review: The vaccine supply chain. In: *European Journal of Operational Research* **268**(1), pp. 174–192.

Duijzer, L., Van Jaarsveld, W., Wallinga, J. and Dekker, R. (2018b). Dose-optimal vaccine allocation over multiple populations. In: *Production and operations management* **27**(1), pp. 143–159.

El Ghaoui, L., Oustry, F. and Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. In: *SIAM journal on optimization* **9**(1), pp. 33–52.

El Hassan, N., Maddah, B. and Abdelaziz, F. B. (2022). "Application of Optimal Stopping Theory to Pandemic Lockdown Policies". In: *2022 International Conference on Decision Aid Sciences and Applications (DASA)*. IEEE, pp. 779–783.

Epstein, S. K., Ciubotaru, R. L. and Wong, J. B. (1997). Effect of failed extubation on the outcome of mechanical ventilation. In: *Chest* **112**(1), pp. 186–192.

Erat, S. and Kavadias, S. (2008). Sequential testing of product designs: implications for learning. In: *Management Science* **54**(5), pp. 956–968.

ESCAP, U., UPU, U. P. U. and World Health Organization (WHO and others) (2020). *How COVID-19 is Changing the World: A Statistical Perspective*. `https://data.unicef.org/resources/how-covid-19-is-changing-the-world-a-statistical-perspective/`.

Esteban-Pérez, A. and Morales, J. M. (2022). Distributionally robust stochastic programs with side information based on trimmings. In: *Mathematical Programming* **195**, pp. 1069–1105.

Feng, Y. and Gallego, G. (1995). Optimal starting times for end-of-season sales and optimal stopping times for promotional fares. In: *Management science* **41**(8), pp. 1371–1391.

Filar, J. A., Kallenberg, L. C. and Lee, H.-M. (1989). Variance-penalized Markov decision processes. In: *Mathematics of Operations Research* **14**(1), pp. 147–161.

Filar, J. A., Krass, D. and Ross, K. W. (1995). Percentile performance criteria for limiting average Markov decision processes. In: *IEEE Transactions on Automatic Control* **40**(1), pp. 2–10.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. In: *Finance and stochastics* **6**(4), pp. 429–447.

Fu, C., Zhou, M., Lou, Z., Xie, J., Sim, M. and Tan, K. B. (2022). Analytics with Robust Epidemiological Compartmental Optimization Models. In: *Major Revision at Manufacturing & Service Operations Management. Available at SSRN 3869521*.

Goldberg, D. A. and Chen, Y. (2018). Beating the curse of dimensionality in options pricing and optimal stopping. In: *arXiv preprint arXiv:1807.02227*.

Greenland, S. (2007). Bayesian perspectives for epidemiological research. II. Regression analysis. In: *International journal of epidemiology* **36**(1), pp. 195–202.

Grefenstette, J. J., Brown, S. T., Rosenfeld, R., DePasse, J., Stone, N. T., Cooley, P. C., Wheaton, W. D., Fyshe, A., Galloway, D. D., Sriram, A., et al. (2013). FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. In: *BMC public health* **13**(1), pp. 1–14.

Gu, Y. (2020). *Projections using machine learning. Retrieved October 10.* `https://covid19-projections.com/`.

Halpern, N. A. and Pastores, S. M. (2015). Critical care medicine beds, use, occupancy and costs in the United States: a methodological review. In: *Critical care medicine* **43**(11), p. 2452.

Hau, J. L., Petrik, M. and Ghavamzadeh, M. (2023). "Entropic risk optimization in discounted MDPs". In: *International Conference on Artificial Intelligence and Statistics*. PMLR, pp. 47–76.

Hong, L. J., Huang, Z. and Lam, H. (2021). Learning-based robust optimization: Procedures and statistical guarantees. In: *Management Science* **67**(6), pp. 3447–3467.

Howard, R. A. and Matheson, J. E. (1972). Risk-sensitive Markov decision processes. In: *Management science* **18**(7), pp. 356–369.

Isermann, H. (1982). Linear lexicographic optimization. In: *Operations-Research-Spektrum* **4**(4), pp. 223–228.

Jelito, D., Pitera, M. and Stettner, Ł. (2021). Risk sensitive optimal stopping. In: *Stochastic Processes and their Applications* **136**, pp. 125–144.

Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A. and Mark, R. (2020). Mimic-iv. In: *PhysioNet. Available online at: https://physionet. org/content/mimiciv/1.0/(accessed August 23, 2021)*, pp. 49–55.

Joynt, K. E., Jha, A. K., et al. (2012). Thirty-day readmissions—truth and consequences. In: *New England Journal of Medicine* **366**(15), pp. 1366–1369.

Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. In: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* **115**(772), pp. 700–721.

Kosmala, T. and Moriarty, J. (2023). The risk-sensitive optimal stopping problem: geometric solution and algorithms. In: *arXiv preprint arXiv:2306.17623*.

Kotas, J. (2019). Optimal stopping for response-guided dosing. In: *Networks and Heterogeneous Media* **14**(1), pp. 43–52.

Kozakiewicz, W. (1947). On the convergence of sequences of moment generating functions. In: *The Annals of Mathematical Statistics*, pp. 61–69.

Krishnamurthy, V. (2016). *Partially observed Markov decision processes: from filtering to controlled sensing*. Cambridge University Press.

Krishnamurthy, V., Aprem, A. and Bhatt, S. (Sept. 2018). Multiple stopping time POMDPs: structural results & application in interactive advertising on social media. In: *Automatica* **95**, pp. 385–398.

Lauri, M., Hsu, D. and Pajarinen, J. (2022). Partially observable markov decision processes in robotics: A survey. In: *IEEE Transactions on Robotics* **39**(1), pp. 21–40.

Lekone, P. E. and Finkenstädt, B. F. (2006). Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. In: *Biometrics* **62**(4), pp. 1170–1177.

Lewnard, J. A., Mbah, M. L. N., Alfaro-Murillo, J. A., Altice, F. L., Bawo, L., Nyenswah, T. G. and Galvani, A. P. (2014). Dynamics and control of Ebola virus transmission in Montserrado, Liberia: a mathematical modelling analysis. In: *The Lancet Infectious Diseases* **14**(12), pp. 1189–1195.

Li, M. L., Bouardi, H. T., Lami, O. S., Trikalinos, T. A., Trichakis, N. and Bertsimas, D. (2023). Forecasting COVID-19 and analyzing the effect of government interventions. In: *Operations Research* **71**(1), pp. 184–201.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W. and Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). In: *Science* **368**(6490), pp. 489–493.

Li, X., Zhong, H. and Brandeau, M. L. (2022). Quantile Markov decision processes. In: *Operations research* **70**(3), pp. 1428–1447.

Li, X., Zhao, Z., Zhu, X. and Wyatt, T. (2011). Covering models and optimization techniques for emergency response facility location and planning: a review. In: *Mathematical Methods of Operations Research* **74**, pp. 281–310.

Liu, F., Chen, Z. and Wang, S. (2023). Globalized distributionally robust counterpart. In: *INFORMS Journal on Computing* **35**(5), pp. 1120–1142.

Liu, Q., Jiang, D., Shi, N., Hayat, T. and Ahmad, B. (2017). Stationary distribution and extinction of a stochastic SEIR epidemic model with standard incidence. In: *Physica A: Statistical Mechanics and its Applications* **476**, pp. 58–69.

Long, D. Z., Sim, M. and Zhou, M. (2023). Robust satisficing. In: *Operations Research* **71**(1), pp. 61–82.

Long, E. F., Nohdurft, E. and Spinler, S. (2018). Spatial resource allocation for emerging epidemics: A comparison of greedy, myopic, and dynamic policies. In: *Manufacturing & Service Operations Management* **20**(2), pp. 181–198.

Ma, S., Ning, S. and Yang, S. (2022). COVID-19 and Influenza Joint Forecasts Using Internet Search Information in the United States. In: *arXiv preprint arXiv:2202.02621*.

Mak, H.-Y., Dai, T. and Tang, C. S. (2022). Managing two-dose COVID-19 vaccine rollouts with limited supply: Operations strategies for distributing time-sensitive resources. In: *Production and operations management* **31**(12), pp. 4424–4442.

Marcus, S. I., Fernández-Gaucherand, E., Hernández-Hernandez, D., Coraluppi, S. and Fard, P. (1997). "Risk sensitive Markov decision processes". In: *Systems and control in the twenty-first century*. Springer, pp. 263–279.

Marianov, V. and ReVelle, C. (1996). The queueing maximal availability location problem: a model for the siting of emergency vehicles. In: *European Journal of Operational Research* **93**(1), pp. 110–120.

McConville, J. F. and Kress, J. P. (2012). Weaning patients from the ventilator. In: *New England Journal of Medicine* **367**(23), pp. 2233–2239.

Mittal, A. and Hanasusanto, G. A. (2022). Finding minimum volume circumscribing ellipsoids using generalized copositive programming. In: *Operations Research* **70**(5), pp. 2867–2882.

Moharrami, M., Murthy, Y., Roy, A. and Srikant, R. (2022). A Policy Gradient Algorithm for the Risk-Sensitive Exponential Cost MDP. In: *arXiv preprint arXiv:2202.04157*.

Moran, J. L., Bristow, P., Solomon, P. J., George, C. and Hart, G. K. (2008). Mortality and length-of-stay outcomes, 1993–2003, in the binational Australian and New Zealand intensive care adult patient database. In: *Critical care medicine* **36**(1), pp. 46–61.

Nasrollahzadeh, A. A. and Khademi, A. (2020). Optimal stopping of adaptive dose-finding trials. In: *Service Science* **12**(2-3), pp. 80–99.

Peng, C., Delage, E. and Li, J. (2020). Probabilistic envelope constrained multiperiod stochastic emergency medical services location model and decomposition scheme. In: *Transportation Science* **54**(6), pp. 1471–1494.

Qi, J. (2017). Mitigating delays and unfairness in appointment systems. In: *Management Science* **63**(2), pp. 566–583.

Rohatgi, V. and Saleh, E. (2015). *An introduction to probability and statistics*. John Wiley & Sons.

Ross, R. (1916). An application of the theory of probabilities to the study of a priori pathometry.—Part I. In: *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character* **92**(638), pp. 204–230.

Rousseeuw, P. J. and Hubert, M. (2011). Robust statistics for outlier detection. In: *Wiley interdisciplinary reviews: Data mining and knowledge discovery* **1**(1), pp. 73–79.

Schwartz, B., Ben-Haim, Y. and Dacso, C. (2011). What makes a good decision? Robust satisficing as a normative standard of rational decision making. In: *Journal for the Theory of Social Behaviour* **41**(2), pp. 209–227.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic orders*. Springer.

Shapiro, A., Dentcheva, D. and Ruszczynski, A. (2009). *Lectures on stochastic programming: modeling and theory*. SIAM.

Shi, P., Helm, J. E., Deglise-Hawkinson, J. and Pan, J. (2021). Timing it right: Balancing inpatient congestion vs. readmission risk at discharge. In: *Operations Research* **69**(6), pp. 1842–1865.

Shrank, W. H., Rogstad, T. L. and Parekh, N. (2019). Waste in the US health care system: estimated costs and potential for savings. In: *Jama* **322**(15), pp. 1501–1509.

Sim, M., Tang, Q., Zhou, M. and Zhu, T. (2021). The analytics of robust satisficing: predict, optimize, satisfice, then fortify. In: *SSRN 3829562*.

Smith, J. E. and Winkler, R. L. (2006). The optimizer's curse: skepticism and postdecision surprise in decision analysis. In: *Management Science* **52**(3), pp. 311–322.

Sturt, B. (2023). A nonparametric algorithm for optimal stopping based on robust optimization. In: *Operations Research.*

Sun, L., DePuy, G. W. and Evans, G. W. (2014). Multi-objective optimization models for patient allocation during a pandemic influenza outbreak. In: *Computers & Operations Research* **51**, pp. 350–359.

Sun, P. and Freund, R. (2004). Computation of minimum-volume covering ellipsoids. In: *Operations Research* **52**(5), pp. 690–706.

Tanner, M. W. and Ntaimo, L. (2010). IIS branch-and-cut for joint chance-constrained stochastic programs and application to optimal vaccine allocation. In: *European Journal of Operational Research* **207**(1), pp. 290–296.

Thille, A. W., Richard, J.-C. M. and Brochard, L. (June 2013). The Decision to Extubate in the Intensive Care Unit. In: *American Journal of Respiratory and Critical Care Medicine* **187**(12), pp. 1294–1302.

Todd, M. J. (2016). *Minimum-volume ellipsoids: Theory and algorithms.* SIAM.

Van Aelst, S. and Rousseeuw, P. (2009). Minimum volume ellipsoid. In: *Wiley Interdisciplinary Reviews: Computational Statistics* **1**(1), pp. 71–82.

Vincent, J.-L., Moreno, R., Takala, J., Willatts, S., De Mendonça, A., Bruining, H., Reinhart, C., Suter, P. and Thijs, L. G. (1996). *The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure.*

Wang, I., Becker, C., Van Parys, B. and Stellato, B. (2023). Learning Decision-Focused Uncertainty Sets in Robust Optimization. In: *arXiv preprint arXiv:2305.19225.*

Wilmore, D. W. and Kehlet, H. (2001). Management of patients in fast track surgery. In: *British Medical Journal* **322**(7284), pp. 473–476.

Wong, W.-T., Lai, V. K., Chee, Y. E. and Lee, A. (2016). Fast-track cardiac care for adult cardiac surgical patients. In: *Cochrane Database of Systematic Reviews* (9).

Xia, L. (2020). Risk-Sensitive Markov Decision Processes with Combined Metrics of Mean and Variance. In: *Production and Operations Management* **29**(12), pp. 2808–2827.

Xie, J., Loke, G. G., Sim, M. and Lam, S. W. (2023). The analytics of bed shortages: Coherent metric, prediction, and optimization. In: *Operations Research* **71**(1), pp. 23–46.

Yarmand, H., Ivy, J. S., Denton, B. and Lloyd, A. L. (2014). Optimal two-phase vaccine allocation to geographically different regions under uncertainty. In: *European Journal of Operational Research* **233**(1), pp. 208–219.

Zhang, X. and Wang, K. (2014). Stochastic SEIR model with jumps. In: *Applied Mathematics and Computation* **239**, pp. 133–143.

Zhou, M., Loke, G. G., Bandi, C., Liau, Z. Q. G. and Wang, W. (2022a). Intraday scheduling with patient re-entries and variability in behaviours. In: *Manufacturing & Service Operations Management* **24**(1), pp. 561–579.

Zhou, M., Sim, M. and Lam, S.-W. (2022b). Advance admission scheduling via resource satisficing. In: *Production and Operations Management* **31**(11), pp. 4002–4020.

Zhu, F., Lee, A. and Chee, Y. E. (2012). Fast-track cardiac care for adult cardiac surgical patients. In: *Cochrane Database of Systematic Reviews* (10).

Zhu, T., Xie, J. and Sim, M. (2022). Joint estimation and robustness optimization. In: *Management Science* **68**(3), pp. 1659–1677.

Zhu, T., Yeo, N. T. B., Gao, S. Y. and Loke, G. G. (2023). Inventory-Responsive Donor-Management Policy: A Tandem Queueing Network Model. In: *Manufacturing & Service Operations Management.*

Zuckerman, R. B., Sheingold, S. H., Orav, E. J., Ruhter, J. and Epstein, A. M. (2016). Readmissions, observation, and the hospital readmissions reduction program. In: *New England Journal of Medicine* **374**(16), pp. 1543–1551.