Algorithms for Learning and Incentive Design in Societal-Scale Systems

Kshitij Kulkarni

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2024-221
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2024/EECS-2024-221.html

December 19, 2024

Algorithms for Learning and Incentive Design in Societal-Scale Systems

By

Kshitij Kulkarni

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor S. Shankar Sastry, Chair
Professor Alexandre Bayen
Professor Anil Aswani
Professor Wesley Holliday

Fall 2024

Algorithms for Learning and Incentive Design in Societal-Scale Systems

Abstract

Algorithms for Learning and Incentive Design in Societal-Scale Systems

by

Kshitij Kulkarni

Doctor of Philosophy in Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor S. Shankar Sastry, Chair

This thesis studies problems that decision-makers face in environments that react or adapt to them. Chapter 2 studies incentive design problems where a decision-maker sets incentives over strategic agents who adapt over time and introduces a two-timescale algorithm for solving these problems. Applications include atomic networked aggregative games and non-atomic routing games. Chapter 3 considers the problem of inferring causal relationships between the state of a dynamical system and the occurrence of a rare event associated with the system and introduces an algorithm for solving this problem by aggregating data across time. Chapter 4 studies the formation of coalitions of charging stations in electric vehicle charging games and analyzes their performance relative to regimes where there are no coalitions and where there are coalitions of maximal size. Chapter 5 studies the redesign of congestion pricing schemes in settings where there are heterogeneities between the values-of-time of travelers during traffic routing. Chapter 6 looks at group decision-making problems in dynamic settings, where the preferences of agents might change, in contrast to the classical setting. In this setting, we prove a representation theorem that characterizes long-run policies.

Contents

Contents **i**

1 Introduction **1**

2 Adaptive Incentive Design with Learning Agents **4**
2.1 Introduction 4
2.2 Related Works 6
2.3 Model 8
2.4 General results 11
2.5 Applications 18
2.6 Discussion and Future Work 25

3 Causal Discovery for Rare Events **26**
3.1 Introduction 26
3.2 Related Works 27
3.3 Preliminaries 28
3.4 Methods 31
3.5 Results 33
3.6 Discussion and Future Work 36

4 Understanding Coalitions Between EV Charging Stations **39**
4.1 Introduction 39
4.2 Related Works 41
4.3 Model 42
4.4 Results 43
4.5 Numerical Results 51
4.6 Discussion and Future Work 52

5 Redesigning Congestion Pricing **54**
5.1 Introduction 54
5.2 Related Works 58
5.3 Model 59
5.4 Computation Methods 62

5.5 Model Calibration for the San Francisco Bay Area 69
5.6 Analysis 74
5.7 Discussion 78

6 Social Choice with Changing Preferences **82**
6.1 Introduction 82
6.2 Social Choice Theory 83
6.3 Dynamic Decision-Making for Groups 85
6.4 Quasi-Utilitarian Characterization 87
6.5 Discussion and Future Work 89

7 Conclusion **91**

Bibliography **92**

A Proofs for Chapter 2 **107**

B Proofs for Chapter 3 and Additional Experiments **112**

C Proofs for Chapter 4 **118**

D Proofs for Chapter 6 **121**

Acknowledgments

I would like to thank my advisor, Prof. Shankar Sastry, for his guidance throughout the years. From Shankar, I learned how to think both rigorously and broadly about problems. Shankar's relentless recommendations to check dimensions, make sure notation is correct, and all mathematical objects are well-defined convinced me of the following truth: sweating the details matters. At the same time, he insisted that I not focus on little problems, but rather think hard about the big stuff, teaching me another truth: what you choose to work on matters.

I would also like to thank my committee (Prof. Alexandre Bayen, Prof. Anil Aswani, and Prof. Wesley Holliday) for their advice and time, as well as the co-authors of the papers contained within this thesis (Chinmay Maheshwari, Chih-Yuan Chiu, Prof. Manxi Wu, Druv Pai, Jiarui Yang, Sukanya Kudva, and Prof. Sven Neth).

A big thank you to friends and collaborators who introduced me to the crazy, fun world of crypto: Tarun Chitra, Guillermo Angeris, Theo Diamandis, Prof. Matheus V.X. Ferreira, and Prof. Mallesh Pai.

A word about my intellectual development throughout graduate school: writings by René Girard, Robert Nozick, and Peter Thiel greatly influenced my understanding of how technology relates to people. We live in a time when technology and politics are in a deadly race for the future. On one hand is a boundless frontier, full of freedom, where people can do new things and try to explore the universe. On the other is stagnation, or worse.

Finally, thank you to my family and God.

Chapter 1

Introduction

We are increasingly surrounded by autonomous systems that learn from and make decisions over the actions of strategic agents that they interact with. Some instances where such systems are prevalent include traffic routing, electric vehicle charging, and the Internet. In these applications, it is important to combine modern computing technologies with incentives to improve the outcomes for all participants. This thesis studies several problems related to the design and analysis of such systems. We describe the chapters of the thesis in more detail:

Adaptive incentive design with learning agents. Incentive mechanisms are commonly used to improve the outcome of the interaction of a large number of strategic agents. Operators often face adaptive agents who repeatedly update their strategies in response to the incentive mechanism. We propose an adaptive incentive mechanism that updates incentives based on agents who repeatedly update their behavior as part of a learning process; the incentive mechanism runs at a timescale that is much slower than the agents' updates. This results in a *two-timescale* coupled dynamical system. The incentive mechanism uses the externality incurred by the agents' strategies, relative to a cost specified by the operator. We provide sufficient conditions on the underlying game between the agents such that the incentive mechanism and agents' updates jointly converge to a socially optimal point. To demonstrate the practical applicability of this mechanism, we apply the theory to atomic networked aggregative games and to non-atomic routing games. This chapter is adapted from [132] and is joint work with Chinmay Maheshwari, Prof. Manxi Wu, and Prof. Shankar Sastry.

Causal discovery for rare events. Causal phenomena associated with rare events occur across a wide range of engineering problems, such as risk-sensitive safety analysis, accident analysis and prevention, and extreme value theory. However, current methods for causal discovery are often unable to uncover causal links between random variables in a dynamic setting that manifest only when the variables first experience low-probability realizations. To address this issue, we introduce a novel statistical independence test on data collected from time-invariant dynamical systems in which rare but consequential events occur. In particular, we exploit the time-invariance of the underlying data to construct a superimposed dataset of the system state before rare events happen at different timesteps. We then design a conditional independence test on the reorganized data. We provide sample complexity bounds for the consistency of our method, and validate its performance across various simulated and real-world datasets, including incident data collected from the Caltrans Performance Measurement System (PeMS). This chapter is adapted from [40] and is joint work with Chih-Yuan Chiu and Prof. Shankar Sastry.

Understanding coalitions between EV charging stations. The rapid growth of electric vehicles (EVs) is driving the expansion of charging infrastructure globally. As charging stations become ubiquitous, their substantial electricity consumption can influence grid operation and electricity pricing. Naturally, some groups of charging stations, which could be jointly operated by a company, may coordinate to decide their charging profile. While coordination among all charging stations is ideal, it is unclear if coordination of some charging stations is better than no coordination. We analyze this intermediate regime between no and full coordination of charging stations. We model EV charging as a non-cooperative aggregative game, where each station's cost is determined by both monetary payments tied to reactive electricity prices on the grid and its sensitivity to deviations from a desired charging profile. We consider a solution concept that we call $\mathcal{C}$-Nash equilibrium, which is tied to a coalition $\mathcal{C}$ of charging stations coordinating to reduce their costs. We provide sufficient conditions, in terms of the demand and sensitivity of charging stations, to determine when independent and uncoordinated operation of charging stations could result in lower overall costs to charging stations, the coalition, and charging stations outside the coalition. Somewhat counter to common intuition, we show numerical instances where allowing charging stations to operate independently is better than coordinating a subset of stations as a coalition. Jointly, these results provide operators of charging stations insights into how to coordinate their charging behavior, and open several research directions. This chapter is adapted from [106] and is joint work with Sukanya Kudva, Chinmay Maheshwari, Prof. Anil Aswani, and Prof. Shankar Sastry.

Redesigning congestion pricing. We study congestion pricing schemes that not only minimize total travel time, but also incorporate the heterogeneous values-of-time of travelers via a distribution-welfare objective. Our analysis builds on a congestion game model with heterogeneous traveler populations. We present four pricing schemes that account for practical considerations, such as the ability to charge differentiated tolls to various traveler populations and the option to toll all or only a subset of edges in the network. We evaluate our pricing schemes in a calibrated freeway network of the San Francisco Bay Area. We demonstrate that the proposed congestion pricing schemes improve both the total travel time and the distribution-welfare objective compared to the current pricing scheme. Our results further show that pricing schemes charging differentiated prices to traveler populations with varying values-of-time lead to a less variable distribution of travel costs compared to those that charge a homogeneous price to all. This chapter is adapted from [129], and is joint work with Chinmay Maheshwari, Druv Pai, Jiarui Yang, Prof. Manxi Wu, and Prof. Shankar Sastry.

Social choice with changing preferences. We study group decision making with changing preferences as a Markov Decision Process (MDP). We are motivated by the increasing prevalence of automated decision-making systems when making choices for groups of people over time. Our main contribution is to show how classic representation theorems from social choice theory can be adapted to characterize optimal policies in this dynamic setting. We provide an axiomatic characterization of MDP reward functions that agree with the Utilitarianism social welfare functionals of social choice theory. We also provide discussion of cases when the implementation of social choice-theoretic axioms may fail to lead to long-run optimal outcomes. This chapter is adapted from [107] and is joint work with Prof. Sven Neth.

Chapter 2

Adaptive Incentive Design with Learning Agents

> Best way to predict the future:
> just follow the incentives.
>
> Vivek Ramaswamy

This chapter can be found in [132].

2.1 Introduction

Incentive mechanisms play a crucial role in many societal systems, where outcomes are governed by the interaction of a large number of self-interested users (or algorithms on their behalf) and a system operator. The outcome of such strategic interactions, characterized by the Nash equilibrium, is often sub-optimal because of the fact that individual agents often do not account for the *externality* of their actions (i.e., how their actions affect the costs of others), when minimizing their own costs. An important way to address this sub-optimality is to provide agents with incentives that align their individual goal of cost minimization with the goal of minimizing the total cost of the system [120, 94, 21].

The system operator faces two key challenges when designing such incentive mechanisms. First, due to the system's scale and privacy concerns, the designer may only have access to sampled information about the cost of agents in the system. Second, the system operator often needs to consider the adaptive behavior of agents, who repeatedly update their strategies in response to the incentive mechanism, especially when the physical system experiences a random shock and agents are trying to reach a new equilibrium [19, 43, 133]. As a result, the designer cannot infer the equilibrium to predict agents' behavior, making incentive design a challenging task. This leads to the following question:

> **Main Question (Q):** How can a system operator learn an incentive mechanism that achieves social optimality based on limited or sampled information about the behavior of agents who are adaptively updating their strategies?

To answer this question, we propose an adaptive incentive mechanism that updates incentives based on the behavior of agents, which is repeatedly updated as part of a learning process, and the gradient information of their cost functions. This results in a *coupled* dynamical system comprising incentive and strategy updates. Our proposed framework can be applied to both atomic and non-atomic games.

Our proposed incentive mechanism has three key features. First, the incentive update incorporates the *externality* incurred by the agents' current strategy, which is evaluated as the difference between their own marginal cost and the marginal cost for the entire system. This feature guarantees that any fixed point of the coupled learning dynamics induces a socially optimal outcome: the incentive provided to each agent equals their externality given their equilibrium strategy, and the associated Nash equilibrium is a socially optimal strategy (Proposition 2). Second, the incentive mechanism only requires *limited information* about the behavior of agents. Specifically, it is agnostic to the strategy update dynamics used by agents and only requires information about the gradient of the cost function of agents at the current strategy to evaluate the externality (Remark 2.3.2). Third, the incentive update occurs at a *slower timescale* compared to the agents' strategy updates. This slower evolution of incentives is a desirable characteristic because frequent incentive updates may hamper agents' participation. It allows agents to consider the incentives as static while updating their strategies.

We provide sufficient conditions on the underlying game such that the set of fixed points is a singleton, making the socially optimal incentive mechanism unique (Proposition 2). We also characterize sufficient conditions that guarantee the convergence of the coupled dynamical system (Proposition 3). Since the convergent strategy profile and incentive mechanism correspond to a fixed point that is also socially optimal, these sufficient conditions ensure that the dynamic incentive mechanism eventually induces a socially optimal outcome in the long run. To prove this result, we leverage the timescale separation between the strategy update and the incentive update, using results from the theory of two-timescale dynamical systems [30]. This theory allows us to analyze the convergence of the coupled dynamical system in two steps by decoupling the convergence of the strategy update and the incentive mechanism update. First, the convergence of the strategy update (on the faster timescale) is analyzed by viewing the incentive mechanism (on the slower timescale) as static. In our work, we off-load this convergence to the rich literature on learning in games (e.g., [74, 170, 142]). Second, the convergence of the incentive mechanism update is analyzed via the associated continuous-time dynamical system, evaluated at the converged value of the strategy update (i.e., Nash equilibrium). The sufficient conditions developed in Proposition 3 are built on results on the convergence of non-linear dynamical systems from cooperative dynamical systems theory [93] and Lyapunov-based methods [171].

To demonstrate the usefulness of the adaptive incentive mechanism, we apply the theory to two classes of practically relevant games: (i) atomic quadratic networked aggregative games and (ii) non-atomic routing games. In atomic networked quadratic games, each agent's cost function depends on their own strategy as well as the aggregate strategies of their neighbors. This aggregation is done through a linear combination of the neighboring agents' strategies with weights characterized by a network matrix. Our proposed incentive mechanism enables the system operator to adaptively update incentives, based on the externality of each agent's strategy on their neighbors, while the agents are learning the equilibrium strategies. Our results when specialized to this setup provide sufficient conditions on the network matrix to ensure that the system converges to a socially optimal outcome.

Furthermore, in non-atomic routing games, agents (travelers) make routing decisions in a congestible network with multiple origin-destination pairs. The system operator imposes incentives in the form of toll prices on each network edge. Our proposed incentive mechanism adaptively gets updated using only the observed flow on edges and the gradient of the edge latency function. Agents can adopt a variety of strategy updates that lead to the equilibrium of the routing game. We show that the adaptive incentive mechanism converges to the toll price that minimizes total congestion.

The chapter is organized as follows: in Section 2.3, we describe the setup of both the atomic and non-atomic games considered here. In addition, we also introduce the joint strategy and incentive update considered in this paper. We present the main results on the fixed point and convergence of the adaptive incentive mechanism in Section 3.5. In Section 2.5, we study the application of our proposed adaptive incentive mechanism on atomic networked aggregative games (Section 2.5) and non-atomic routing games (Section 2.5). We conclude with a discussion in Section 2.6.

2.2 Related Works

Two-timescale Learning Dynamics

Learning dynamics where incentives are updated at a slower timescale than the strategy of agents have been studied in [141, 38, 164, 152, 167, 126, 121]. Specifically, [141] studies Stackelberg games with a single leader and a population of followers. In their approach, the leader deploys a gradient-based update and followers update their strategies using replicator dynamics. [38] and [164] consider incentive design in affine congestion games, where the incentives are updated using a distributed version of gradient descent. [152] studies incentive design to control traffic congestion on a single highway using gradient based incentive updates. [167] studies the problem of incentive design while learning the cost functions of agents. The authors assume that both the cost functions and incentive policies are linearly parameterized, and the incentive updates rely on the knowledge of agents' strategy update rules instead of just the current strategy as in our setting. Additionally, [126, 121] propose a two-timescale discrete-time learning dynamics where agents adopt a mirror descent-based strategy update and the system operator updates the incentive parameter according to a gradient-based method. The convergence of these gradient-based learning dynamics relies on the assumption that the social cost, given agents' equilibrium strategy, is convex in the incentive parameter. However, the convexity assumption is restrictive, as highlighted by an example in Appendix A.

Single-timescale Learning Dynamics

The problem of *steering* non-cooperative agents to a desired Nash equilibrium by considering an incentive update which runs on the same timescale as strategy updates is studied in [174, 175, 194]. In [174], the authors study such updates in the specific setting of quadratic aggregative games. In [175], the authors consider a setup where the agents' costs are only dependent on their own action and the price signal which is provided by operator which is used by agents based on their individual reliability over the operator. The aim is to design this price signal so that the reliability on the price signal is retained while steering the equilibrium to the desired point. In [194], the authors consider the problem of steering no-regret learning agents to an optimal equilibrium, but computing the incentive mechanism requires solving an optimization problem at each time step.

Learning in Stackelberg Games

Our work is also related to the literature on learning in Stackelberg games, where the operator often has limited information about the game between agents and needs to design the optimal mechanism by dynamically receiving feedback from agents' responses (for instance, refer to [28, 119, 158, 16, 47, 98, 134]). Typically, this line of work posits structural assumptions on the game between followers (e.g., a finite action space of agents, or linearly parametrized utility functions of agents)[28, 119, 158, 16, 47, 98], or only ensures convergence to a locally optimal solution [134].

In comparison to the three former lines of research, our externality-based incentive design is applicable for a broad class of games (with continuous action spaces and potentially non-linear utility functions), while providing sufficient conditions to ensure asymptotic convergence to the optimal solution.

Notation

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we use $\nabla_{x_i} f(x)$ to denote $\frac{\partial f}{\partial x_i}(x)$, the derivative of f with respect to x_i for any $i \in \{1, 2, ..., n\}$ and ∇f to be the Jacobian of the function. For any set A we use $\mathsf{conv}(A)$ to denote the convex hull of the set. For any set $X \subseteq \mathbb{R}^n$, we say that a function $f : X \to \mathbb{R}$ is *Lipschitz* if there exists a positive scalar L such that $\|f(x) - f(x')\| \leq L\|x - x'\|$ for every $x, x' \in X$. For any vector $x \in X$ and any positive scalar $r > 0$, the set $\mathcal{B}_r(x) = \{x' \in X | \|x' - x\| < r\}$ denotes the r-radius neighborhood of the vector x. For any set X, we define $\mathsf{boundary}(X)$ and $\mathsf{int}(X)$ to be the boundary and interior of the set X, respectively. Finally, for any function $f(\cdot)$, we denote the domain of the function by $\mathsf{dom}(f)$. For any vector $x \in \mathbb{R}^n$, we define $\mathsf{diag}(x) \in \mathbb{R}^{n\times n}$ to be a diagonal matrix with diagonal entries corresponding to entries of x.

2.3 Model

We now introduce both atomic and non-atomic static games, and present the adaptive incentive design approach proposed in this work.

Static Games

Atomic Games

Consider a game G with a finite set of players $\mathcal{I}$. The strategy of each player $i \in \mathcal{I}$ is $x_i \in X_i$, where X_i is a non-empty and closed interval in $\mathbb{R}$. The joint strategy profile of all players is $x = (x_i)_{i\in\mathcal{I}}$, and the set of all joint strategy profiles is $X \coloneqq \prod_{i\in\mathcal{I}} X_i$. The cost function of each player $i \in \mathcal{I}$ is $\ell_i : \mathbb{R}^{|\mathcal{I}|} \to \mathbb{R}$. We measure the outcome of our games by costs instead of utilities. Equivalently, the utility of each player is the negative value of the cost.

A *system operator* designs incentives by setting a payment $p_i x_i \in \mathbb{R}$ for each player i that is linear in their strategy x_i. Here, $p_i \in \mathbb{R}$ represents the marginal payment for every unit increase in strategy of player i. The value of $p_i x_i$ can either be negative or positive, which represents a marginal subsidy or a marginal tax, respectively. Given the incentive mechanism $p = (p_i)_{i\in I}$, the total cost of each player $i \in \mathcal{I}$ is:

$$c_i(x, p) = \ell_i(x) + p_i x_i, \quad \forall\ x \in X. \tag{2.1}$$

A strategy profile $x^*(p) \in X$ is a *Nash equilibrium* in the atomic game G with the incentive mechanism p if

$$c_i(x_i^*(p), x_{-i}^*(p), p) \le c_i(x_i, x_{-i}^*(p), p),\ \forall\ x_i \in X_i,\ \forall i \in \mathcal{I}.$$

A strategy profile $x^\dagger \in X$ is *socially optimal* if it minimizes the social cost function $\Phi : \mathbb{R}^{|\mathcal{I}|} \to \mathbb{R}$ over X.

Assumption 2.3.1. *The cost function $\ell_i(x_i, x_{-i})$ is strongly convex in x_i for all $i \in \mathcal{I}$ and all $x_{-i} = (x_j)_{j\in\mathcal{I}\setminus\{i\}}$. Moreover, $\ell_i(\cdot)$ is continuously differentiable and its gradient is Lipschitz. Additionally, the social cost function $\Phi(x)$ is twice continuously differentiable, has Lipschitz gradient, and is strictly convex.*

Under Assumption 2.3.1, it is known that, for any $p \in \mathbb{R}^{|\mathcal{I}|}$, the Nash equilibrium $x^*(p)$ exists, is unique and is Lipschitz continuous in p [49]. Moreover, the socially optimal strategy $x^\dagger$ is also unique.

Non-atomic Games

Consider a game $\tilde{G}$ with a finite set of player populations $\tilde{\mathcal{I}}$. Each population $i \in \tilde{\mathcal{I}}$ is comprised of a continuum set of (infinitesimal) players with mass $M_i > 0$. Every (infinitesimal) player in any population can choose an action in a finite set S_i. The strategy distribution of population $i \in \tilde{\mathcal{I}}$ is $\tilde{x}_i = (\tilde{x}_i^j)_{j \in S_i}$, where $\tilde{x}_i^j$ is the mass of players in population i who choose action $j \in S_i$. Then, the set of all strategy distributions of population i is $\tilde{X}_i = \left\{\tilde{x}_i \in \mathbb{R}^{|S_i|} | \sum_{j \in S_i} \tilde{x}_i^j = M_i, \ \tilde{x}_i^j \geq 0, \forall j \in S_i\right\}$. The strategy distribution of all populations is $\tilde{x} = (\tilde{x}_i)_{i \in \tilde{\mathcal{I}}} \in \tilde{X} = \prod_{i \in \tilde{\mathcal{I}}} \tilde{X}_i$. We define $S = \prod_{i \in \tilde{\mathcal{I}}} S_i$. Given a strategy distribution $\tilde{x} \in \tilde{X}$, the cost of players in population $i \in \tilde{\mathcal{I}}$ for choosing action $j \in S_i$ is $\tilde{\ell}_i^j(\tilde{x})$. We denote $\tilde{\ell}_i(\tilde{x}) = (\tilde{\ell}_i^j(\tilde{x}))_{j \in S_i}$ as the vector of costs for each $i \in \tilde{\mathcal{I}}$.

A system operator designs incentives by setting a payment $\tilde{p}_i^j$ for players in population i who choose action $j \in S_i$. Consequently, given the incentive mechanism $\tilde{p} = (\tilde{p}_i^j)_{j \in S_i, i \in \tilde{\mathcal{I}}}$, the total cost experienced by any player, in population $i \in \tilde{\mathcal{I}}$, that chooses action $j \in S_i$ is

$$\tilde{c}_i^j(\tilde{x}, \tilde{p}) = \tilde{\ell}_i^j(\tilde{x}) + \tilde{p}_i^j, \quad \forall \ \tilde{x} \in \tilde{X}. \tag{2.2}$$

A strategy distribution $\tilde{x}^*(\tilde{p}) \in \tilde{X}$ is a Nash equilibrium in the nonatomic game $\tilde{G}$ with incentive $\tilde{p}$ if

$$\begin{aligned} &\tilde{x}_i^{j*}(\tilde{p}) > 0 \\ &\qquad \Rightarrow \tilde{c}_i^j(\tilde{x}^*(\tilde{p}), \tilde{p}) \leq \tilde{c}_i^{j'}(\tilde{x}^*(\tilde{p}), \tilde{p}), \ \forall j, j' \in S_i, \ \forall i \in \tilde{\mathcal{I}}. \end{aligned}$$

A strategy distribution $\tilde{x}^\dagger \in \tilde{X}$ is socially optimal if $\tilde{x}^\dagger$ minimizes the social cost function $\tilde{\Phi} : \mathbb{R}^{|S|} \to \mathbb{R}$.

Assumption 2.3.2. *The function $\tilde{\ell}(\cdot)$ is Lipschitz continuous. Moreover, it is strongly monotone: there exists $\rho > 0$ such that*

$$\left\langle \tilde{\ell}(\tilde{x}) - \tilde{\ell}(\tilde{x}'), \tilde{x} - \tilde{x}' \right\rangle > \rho \|\tilde{x} - \tilde{x}'\|^2, \quad \forall \tilde{x} \neq \tilde{x}' \in \tilde{X}.$$

Additionally, $\tilde{\Phi}(\tilde{x})$ is twice continuously differentiable, has Lipschitz gradient, and strictly convex.

Under Assumption 2.3.2, the Nash equilibrium $x^*(p)$ exists, is unique and Lipschitz continuous for every $p \in \mathbb{R}^{|\tilde{\mathcal{I}}|}$ [170, 49]. Moreover, the socially optimal strategy distribution is also unique.

Coupled Strategy and Incentive Update

We consider a coupled dynamics that jointly updates agents' strategies and the incentive mechanism with discrete timesteps $k \in \mathbb{N}$. At step k, the strategy profile in the atomic game

G (respectively non-atomic game $\tilde{G}$) is $x_k = (x_{i,k})_{i\in\mathcal{I}}$ (respectively $\tilde{x}_k = (\tilde{x}_{i,k})_{i\in\tilde{\mathcal{I}}}$), where $x_{i,k}$ (respectively $\tilde{x}_{i,k}$) is the strategy of player i (population i), and the incentive mechanism is $p_k = (p_{i,k})_{i\in\mathcal{I}}$ (respectively $\tilde{p}_k = (\tilde{p}^j_{i,k})_{j\in S_i, i\in\tilde{\mathcal{I}}}$). The strategy updates and the incentive updates are presented below:

Strategy update.

$$x_{k+1} = (1-\gamma_k)x_k + \gamma_k f(x_k, p_k), \qquad (x\text{-update})$$

$$\tilde{x}_{k+1} = (1-\gamma_k)\tilde{x}_k + \gamma_k \tilde{f}(\tilde{x}_k, \tilde{p}_k). \qquad (\tilde{x}\text{-update})$$

In each step $k+1$, the updated strategy is a linear combination of the strategy in step k (i.e. x_k in G and $\tilde{x}_k$ in $\tilde{G}$), and a new strategy $f(x_k, p_k) \in X$ in G (respectively $\tilde{f}(\tilde{x}_k, \tilde{p}_k) \in \tilde{X}$ in $\tilde{G}$) that depends on the previous strategy and the incentive mechanism. The relative weight in the linear combination is determined by the step-size $\gamma_k \in (0,1)$. We require that for any p (respectively $\tilde{p}$), the fixed point associated with the update (x-update) (respectively ($\tilde{x}$-update)) is a Nash equilibrium, i.e.

$$\begin{aligned} x^*(p) &= \{x : f(x,p) = x\}, \quad \forall p \in \mathbb{R}^{|\mathcal{I}|}, \\ \tilde{x}^*(\tilde{p}) &= \{\tilde{x} : \tilde{f}(\tilde{x},\tilde{p}) = \tilde{x}\}, \quad \forall\, \tilde{p} \in \mathbb{R}^{|\tilde{\mathcal{I}}|}. \end{aligned} \tag{2.3}$$

Some examples of commonly studied learning dynamics (x-update) and ($\tilde{x}$-update) are enumerated below:

1. *Equilibrium update ([47, 98]):* The strategy update incorporates a Nash equilibrium strategy profile with respect to the incentive mechanism in step k:

$$f(x_k, p_k) = x^*(p_k), \quad \text{and } \tilde{f}(\tilde{x}_k, \tilde{p}_k) = \tilde{x}^*(\tilde{p}_k). \tag{2.4}$$

2. *Best response update ([74, 96]):* The strategy update incorporates a best response strategy with respect to the strategy and the incentive mechanism in step k:

$$\begin{aligned} f_i(x_k, p_k) &= \arg\min_{y_i \in X_i} c_i(y_i, x_{-i,k}, p_k), \\ \tilde{f}_i(\tilde{x}_k, \tilde{p}_k) &= \arg\min_{\tilde{y}_i \in \tilde{X}_i} \tilde{y}_i^\top \tilde{c}_i(\tilde{x}_k, \tilde{p}_k). \end{aligned} \tag{2.5}$$

3. *Gradient-based update ([125, 109, 170])*: The strategy update incorporates a *smoothed* strategy update with respect to the strategy and the incentive mechanism in step k:

$$\begin{aligned} f_i(x_k, p_k) &= \arg\min_{y_i \in X_i} z_i(x_k, p_k) y_i - h(y_i), \\ \tilde{f}_i(\tilde{x}_k, \tilde{p}_k) &= \arg\min_{\tilde{y}_i \in \tilde{X}_i} \tilde{c}_i(\tilde{x}_k, \tilde{p}_k) - \tilde{h}(\tilde{y}_i), \end{aligned} \tag{2.6}$$

where $z_i(x_k, p_k) = x_k - \eta D_{x_i} c_i(x_k, p_k)$, η is step size, and $h(\cdot), \tilde{h}(\cdot)$ are regularizers. If $h(\cdot)$ is a quadratic function then the update becomes projected gradient descent [139]. Furthermore, if $\tilde{h}(\cdot)$ is the entropy function then the update becomes a perturbed best-response update [170].

Incentive update.

$$p_{k+1} = (1-\beta_k)p_k + \beta_k e(x_k), \tag{p-update}$$
$$\tilde{p}_{k+1} = (1-\beta_k)\tilde{p}_k + \beta_k \tilde{e}(\tilde{x}_k), \tag{$\tilde{p}$-update}$$

where $e(x) = (e_i(x))_{i\in\mathcal{I}}$, $\tilde{e}(\tilde{x}) = (\tilde{e}_i^j(\tilde{x}))_{j\in S_i, i\in\tilde{\mathcal{I}}}$, and

$$e_i(x) = D_{x_i}\Phi(x) - D_{x_i}\ell_i(x), \quad \forall i \in \mathcal{I}, \tag{2.7a}$$
$$\tilde{e}_i^j(\tilde{x}) = D_{\tilde{x}_i^j}\tilde{\Phi}(\tilde{x}) - \tilde{\ell}_i^j(\tilde{x}), \quad \forall j \in S_i, \quad \forall i \in \tilde{\mathcal{I}}. \tag{2.7b}$$

In (2.7a), $e_i(x)$ is the difference between the marginal social cost, and the marginal cost of player i given x. Similarly, $\tilde{e}_i^j(\tilde{x})$ is the difference between the marginal social cost, and the cost experienced by the players in population i who chooses action j. We refer to $e_i(x)$ and $\tilde{e}_i(\tilde{x}) = (\tilde{e}_i^j(\tilde{x}))_{j\in S_i}$ as the externality of player i and population i, respectively, since they capture the difference in impacts of their strategy on the social cost and individual cost.

The updates (p-update)-($\tilde{p}$-update) modify the incentives on the basis of the externality caused by the players. In each step $k+1$, the updated incentive mechanism is a linear combination of the incentive mechanism in step k (i.e. p_k in G and $\tilde{p}_k$ in $\tilde{G}$), and the externality (i.e. $e(x_k)$ in G and $\tilde{e}(\tilde{x}_k)$ in $\tilde{G}$) given the strategy in step k. The relative weight in the linear combination is determined by the step size $\beta_k \in (0, 1)$.

In summary, the joint evolution of the strategy and the incentive mechanism $(x_k, p_k)_{k=1}^{\infty}$ (respectively $(\tilde{x}_k, \tilde{p}_k)_{k=1}^{\infty}$) in the atomic game G (respectively non-atomic game $\tilde{G}$) is governed by the learning dynamics (x-update)–(p-update) (respectively ($\tilde{x}$-update)–($\tilde{p}$-update)). The step-sizes $(\gamma_k)_{k=1}^{\infty}$ and $(\beta_k)_{k=1}^{\infty}$ determine the speed of strategy updates and incentive updates.

Remark 2.3.1. *Our externality-based incentive update contrasts with the gradient-based incentive update considered in the literature [126]. Gradient-based incentive updates typically require that the equilibrium social cost function $\Phi(x^*(p))$ (respectively $\tilde{\Phi}(\tilde{x}^*(\tilde{p}))$) be strongly convex in the incentive mechanism p (respectively $\tilde{p}$). However, this requirement is too restrictive and is not satisfied even in simple games due to the fact that the equilibrium strategy vector is typically non-convex in the incentive mechanism. In Appendix A, we provide an example of a two-link routing game where the equilibrium social cost function is non-convex in the incentive mechanism, and the gradient-based learning dynamics do not converge.*

Remark 2.3.2. *The incentive update is based on only limited information about the players' behavior. Specifically, it relies solely on the gradients of the social cost and those of the players', evaluated at the players' current strategy profile. Note that it is agnostic to the strategy update dynamics (i.e., (x-update), and ($\tilde{x}$-update)) employed by players.*

2.4 General results

In Section 2.4, we characterize the set of fixed points of the updates (x-update)-(p-update) and ($\tilde{x}$-update)-($\tilde{p}$-update), and show that any fixed point corresponds to a socially optimal incentive mechanism such that the induced Nash equilibrium strategy profile minimizes the social cost. In Section 2.4, we provide a set of sufficient conditions that guarantee the convergence of strategy and incentive updates. Under these conditions, our adaptive incentive mechanism eventually induces a socially optimal outcome.

Fixed point analysis

We first characterize the set of fixed points of the updates (x-update)-(p-update), and ($\tilde{x}$-update)-($\tilde{p}$-update) as follows:

$$\text{Atomic game } G,\ \{(x,p)|f(x,p)=x,\ e(x)=p\}\,, \tag{2.8a}$$

$$\text{Non-atomic game } \tilde{G},\ \left\{(\tilde{x},\tilde{p})|\tilde{f}(\tilde{x},\tilde{p})=\tilde{x},\ \tilde{e}(\tilde{x})=\tilde{p}\right\}. \tag{2.8b}$$

Using (2.3), from (2.8a) – (2.8b), we can write the set of incentive mechanisms at the fixed point, $P^\dagger$ as follows:

$$\begin{aligned}&\text{Atomic game } G,\ P^\dagger=\{(p_i^\dagger)_{i\in\mathcal{I}}|e(x^*(p^\dagger))=p^\dagger\},\\&\text{Non-atomic game } \tilde{G},\ \tilde{P}^\dagger=\{(\tilde{p}_i^\dagger)_{i\in\tilde{\mathcal{I}}}|\tilde{e}(\tilde{x}^*(\tilde{p}^\dagger))=\tilde{p}^\dagger\}.\end{aligned} \tag{2.9}$$

That is, at any fixed point, the incentive of each player is set to be equal to the externality evaluated at their equilibrium strategy profile.

Our first result characterizes conditions under which the fixed point set $P^\dagger$ (respectively $\tilde{P}^\dagger$) is non-empty and singleton in G (respectively $\tilde{G}$). Moreover, given any fixed point incentive parameter $p^\dagger \in P^\dagger$ and $\tilde{p}^\dagger \in \tilde{P}^\dagger$, the corresponding Nash equilibrium is socially optimal.

Proposition 2. *Let Assumptions 2.3.1 hold and the strategy set X in an atomic game G be compact. The set $P^\dagger$ is a non-empty singleton set. The unique $p^\dagger \in P^\dagger$ is socially optimal, i.e. $x^*(p^\dagger) = x^\dagger$.*

Moreover, in a non-atomic game $\tilde{G}$ under Assumptions 2.3.2, $\tilde{P}^\dagger$ is a non-empty singleton set. The unique $\tilde{p}^\dagger \in \tilde{P}^\dagger$ is socially optimal, i.e., $\tilde{x}^*(\tilde{p}^\dagger) = \tilde{x}^\dagger$.

Proof. First, we show that $P^\dagger$ is non-empty. That is, there exists $p^\dagger$ such that $e(x^*(p^\dagger)) = p^\dagger$. Define a function $\theta(p) = e(x^*(p))$. We note that Assumption 2.3.1 guarantees that θ is well defined. Thus the problem is reduced to showing the existence of a fixed point of $\theta(\cdot)$, for which we use the Schauder fixed point theorem [177].

We note that Assumption 2.3.1 ensures that $\theta(p)$ is a continuous function. Now, define $K := \{\theta(p) : p \in \mathbb{R}^{|\mathcal{I}|}\} \subseteq \mathbb{R}^{|\mathcal{I}|}$. We claim that the set K is compact. Indeed, this follows by two observations. First, the externality function $e(\cdot)$ is continuous. Second, the range of the function $x^*(\cdot)$ is X, which is a compact set. These two observations ensure that $\theta(p) = e(x^*(p))$ is a bounded function. Let $\tilde{K} := \mathsf{conv}(K)$ be the convex hull of K, which in turn is also a compact set. Denote the restriction of the function θ on the set $\tilde{K}$ as

$\theta_{|\tilde{K}} : \tilde{K} \to \tilde{K}$ where $\theta_{|\tilde{K}}(p) = \theta(p)$ for all $p \in \tilde{K}$. We note that $\theta_{|\tilde{K}}$ is a continuous function from a convex, compact set to itself and therefore, the Schauder fixed point theorem ensures that there exists $p^\dagger \in \tilde{K}$ such that $p^\dagger = \theta_{|\tilde{K}}(p^\dagger) = \theta(p^\dagger)$. This concludes the proof of existence of $p^\dagger$. A similar line of argument as above can be replicated for a non-atomic game $\tilde{G}$ to show that $\tilde{P}^\dagger$ is non-empty.

Next, we show that the incentive $p^\dagger$ aligns the Nash equilibrium with social optimality (i.e. for any $p^\dagger \in P^\dagger$, $x^*(p^\dagger) = x^\dagger$). Fix $p^\dagger \in P^\dagger$. For every $i \in \mathcal{I}$ we have $p_i^\dagger = e_i(x^*(p^\dagger))$. This implies $D_{x_i}\ell_i(x^*(p^\dagger)) + p_i^\dagger = D_{x_i}\Phi(x^*(p^\dagger))$ for every $i \in \mathcal{I}$. This implies

$$J(x^*(p^\dagger), p^\dagger) = \nabla\Phi(x^*(p^\dagger)), \tag{2.10}$$

where $J(x,p)$ is the *game Jacobian* defined as $J_i(x,p) = D_{x_i}\ell_i(x) + p_i$ for every $i \in \mathcal{I}$. Next, under Assumption 2.3.1, using the variational inequality characterization of Nash equilibrium [65], we know that $x^*(p^\dagger)$ is a Nash equilibrium if and only if

$$\langle J(x^*(p^\dagger), p^\dagger), x - x^*(p^\dagger)\rangle \geq 0, \quad \forall\ x \in X. \tag{2.11}$$

From (2.10) and (2.11) observe that

$$\langle \nabla\Phi(x^*(p^\dagger)), x - x^*(p^\dagger)\rangle \geq 0, \quad \forall\ x \in X. \tag{2.12}$$

Further, from the first order conditions of optimality for social cost function we know that $x^\dagger$ is socially optimal if and only if it satisfies

$$\langle \nabla\Phi(x^\dagger), x - x^\dagger\rangle \geq 0, \quad \forall\ x \in X. \tag{2.13}$$

Comparing (2.12) with (2.13), we note that $x^*(p^\dagger)$ is the minimizer of the social cost function Φ. This implies $x^*(p^\dagger) = x^\dagger$ as $x^\dagger$ is the unique minimizer of the social cost function Φ by Assumption 2.3.1.

Similarly for a non-atomic game $\tilde{G}$, we show that the incentive $\tilde{p}^\dagger$ aligns the Nash equilibrium with social optimality (i.e. for any $\tilde{p}^\dagger \in \tilde{P}^\dagger$, $\tilde{x}^*(\tilde{p}^\dagger) = \tilde{x}^\dagger$). Fix $\tilde{p}^\dagger \in \tilde{P}^\dagger$. For every $j \in \tilde{S}_i$ and $i \in \tilde{\mathcal{I}}$, it holds that $\tilde{p}_i^{j\dagger} = \tilde{e}_i^j(\tilde{x}^*(\tilde{p}^\dagger))$. This implies that

$$\tilde{c}_i^j(\tilde{x}^*(\tilde{p}^\dagger), \tilde{p}^\dagger) = D_{\tilde{x}_i^j}\tilde{\Phi}(\tilde{x}^*(\tilde{p}^\dagger)). \tag{2.14}$$

Under Assumption 2.3.2, using the variational inequality characterization of Nash equilibrium for non-atomic games [170], we know that $\tilde{x}^*(\tilde{p}^\dagger)$ is a Nash equilibrium if and only if

$$\langle \tilde{c}(\tilde{x}^*(\tilde{p}^\dagger), \tilde{p}^\dagger), \tilde{x} - \tilde{x}^*(\tilde{p}^\dagger)\rangle \geq 0 \quad \forall\ \tilde{x} \in \tilde{X}. \tag{2.15}$$

From (2.14) and (2.15), we observe that

$$\langle \nabla\tilde{\Phi}(\tilde{x}^*(\tilde{p}^\dagger)), \tilde{x} - \tilde{x}^*(\tilde{p}^\dagger)\rangle \geq 0 \quad \forall\ \tilde{x} \in \tilde{X}. \tag{2.16}$$

From the first order conditions of optimality for the social cost function, we know that $\tilde{x}^\dagger$ is socially optimal if and only if it satisfies

$$\langle \nabla\tilde{\Phi}(\tilde{x}^\dagger), \tilde{x} - \tilde{x}^\dagger\rangle \geq 0, \quad \forall\ \tilde{x} \in \tilde{X}. \tag{2.17}$$

Comparing (2.16) with (2.17), we note that $\tilde{x}^*(\tilde{p}^\dagger)$ is the minimizer of the social cost function $\tilde{\Phi}$. This implies $\tilde{x}^*(\tilde{p}^\dagger) = \tilde{x}^\dagger$ as $\tilde{x}^\dagger$ is the unique minimizer of the social cost function $\tilde{\Phi}$ by Assumption 2.3.2.

Next, we show that the set $P^\dagger$ is singleton. We prove this via contradiction. Suppose $P^\dagger$ contains two element $p_1^\dagger, p_2^\dagger$, which are shown to align the Nash equilibrium with social optimality. Therefore, it holds that $x^\dagger = x^*(p_1^\dagger) = x^*(p_2^\dagger)$. Moreover, due to (2.9) we know that $p_1^\dagger = e(x^*(p_1^\dagger))$ and $p_2^\dagger = e(x^*(p_2^\dagger))$. Consequently, we have $p_1^\dagger = e(x^\dagger) = p_2^\dagger$. The proof of uniqueness of $\tilde{P}^\dagger$ also follows analogously. □

Remark 2.4.1. *Proposition 2 demonstrates that the externality-based incentive updates* (p-update) *and* ($\tilde{p}$-update) *have two advantages. Firstly, these updates guarantee that any fixed point must achieve social optimality. Secondly, the result indicates that the optimal incentive mechanism takes a simple format: the incentive mechanism is linear in players' strategies in atomic games and is a constant vector in non-atomic games.*

Convergence to optimal incentive mechanism

In this subsection, we provide sufficient conditions for the convergence of the strategy and incentive updates (x-update)-(p-update) and ($\tilde{x}$-update)-($\tilde{p}$-update). In order to study the convergence of these updates we make the following assumption:

Assumption 2.4.1. *The step sizes in* (x-update)-(p-update) *and* ($\tilde{x}$-update)-($\tilde{p}$-update) *satisfy*

(i) $\sum_{k=1}^\infty \gamma_k = \sum_{k=1}^\infty \beta_k = +\infty$, $\sum_{k=1}^\infty \gamma_k^2 + \beta_k^2 < +\infty$.

(ii) $\lim_{k\to\infty} \frac{\beta_k}{\gamma_k} = 0$.

In Assumption 2.4.1, *(i)* is a standard assumption on step-sizes that allows us to analyze the convergence of the discrete-time learning updates via a continuous-time differential equation [30]. Additionally, *(ii)* assumes that the incentive update occurs at a *slower timescale* compared to the strategy update.

In view of Assumption 2.4.1-*(ii)*, it is necessary that the strategy updates converge if the incentives are held fixed. This condition holds true for a variety of updates commonly studied in the *learning in games* literature. For instance, the equilibrium update (compare (2.4)) converges to Nash equilibrium in one step. Similarly, the best response update (compare (2.5)) converges in zero sum games [92], potential games [181], and dominance solvable games [145]. The gradient-based update (compare (2.6)) converges in concave games [168]. Therefore, in this work we do not focus on studying the convergence of strategy updates with fixed incentive mechanisms and we pick any *off-the-shelf* convergent strategy update. Our goal is to characterize conditions under which the coupled strategy and incentive updates (x-update)-(p-update) and ($\tilde{x}$-update)-($\tilde{p}$-update) converge. Therefore, we impose the following assumption which is sufficient to ensure convergence of strategy updates (x-update) and ($\tilde{x}$-update) with fixed incentive mechanisms. We assume that if the incentive mechanism is fixed, then the corresponding continuous-time dynamical system converges to the Nash equilibrium associated with that incentive mechanism.

Assumption 2.4.2. *For any incentive mechanism $p, \tilde{p}$, consider the following continuous time dynamical systems associated with* (x-update) *and* ($\tilde{x}$-update)*, respectively,*

$$\dot{x}(t) = f(x(t), p) - x(t), \qquad (x\text{-dynamics})$$

$$\dot{\tilde{x}}(t) = \tilde{f}(\tilde{x}(t), \tilde{p}) - \tilde{x}(t). \qquad (\tilde{x}\text{-dynamics})$$

The equilibrium $x^*(p)$ (respectively $\tilde{x}^*(\tilde{p})$) is the global asymptotically stable fixed point of the dynamics (x-dynamics) *(respectively* ($\tilde{x}$-dynamics)*).*

Using the standard approximation arguments from [31] it is evident that Assumption 2.4.1 and 2.4.2 are sufficient to ensure convergence of strategy updates (x-update) and ($\tilde{x}$-update) with fixed incentive mechanisms. Against the preceding backdrop, in the following result, we characterize the convergence conditions for the coupled strategy and incentive updates.

Proposition 3. *Consider an atomic game G. Suppose that Assumptions 2.3.1, 2.4.1, 2.4.2 and at least one of the following conditions holds:*

(C1) $\frac{\partial e_i(x^*(p))}{\partial p_j} > 0$ for all $p \in \mathbb{R}^n$ and all $i \neq j$. Additionally, at least one of the following conditions hold:

- *(i) If $e_i(x^*(0)) \geq 0$ for every $i \in \mathcal{I}$ then there exists $p \in \mathbb{R}_+^{|\mathcal{I}|}$ such that $e_i(x^*(p)) - p_i \leq 0$ for every $i \in \mathcal{I}$. Moreover, $x_1 \in X, p_1 \in \mathbb{R}_+^{|\mathcal{I}|}$.*
- *(ii) If $e_i(x^*(0)) \leq 0$ for every $i \in \mathcal{I}$ then there exists $p \in \mathbb{R}_-^{|\mathcal{I}|}$ such that $e_i(x^*(p)) - p_i \geq 0$ for every $i \in \mathcal{I}$. Moreover, $x_1 \in X, p_1 \in \mathbb{R}_-^{|\mathcal{I}|}$.*

(C2) There exists a set $\textsf{dom}(V) \subset \mathbb{R}^{|\mathcal{I}|}$ and a continuously differentiable function $V : \textsf{dom}(V) \to \mathbb{R}_+$ such that $V(p^\dagger) = 0$ and $V(p) > 0$ for all $p \neq p^\dagger$. Moreover:

$$\nabla V(p)^\top (e(x^*(p)) - p) < 0 \quad \forall\, p \neq p^\dagger.$$

Additionally, let $\sup_{k\in\mathbb{N}}(\|x_k\| + \|p_k\|) < +\infty$. Then there exist positive scalars $\bar{r}, \bar{\alpha}, \bar{\beta}, \bar{\gamma}$ such that the discrete-time updates (x-update)-(p-update) *in G when initialized at $p_0 \in \mathcal{B}_{\bar{r}}(p^\dagger)$, $x_0 \in \mathcal{B}_{\bar{r}}(x^*(p_0))$, with step-sizes such that*

$$\sup_{k\in\mathbb{N}} \beta_k/\gamma_k \leq \bar{\alpha}, \quad \sup_{k\in\mathbb{N}} \beta_k \leq \bar{\beta}, \quad \sup_{k\in\mathbb{N}} \gamma_k \leq \bar{\gamma}, \tag{2.18}$$

satisfy[1]

$$\lim_{k\to\infty}(x_k, p_k) = (x^\dagger, p^\dagger). \tag{2.19}$$

The same argument holds for a non-atomic game $\tilde{G}$ under Assumptions 2.3.2, 2.4.1, and 2.4.2.

Proof. Under Assumption 2.4.1, the two-timescale approximation theory [30, 32] suggests that the strategy update (x-update) is a *fast transient* while the incentive update (p-update) is a *slow component*. This separation in timescales allow us to study the convergence of these updates in two stages. First, we study the convergence of fast strategy updates, for every fixed value of the incentive. Second, we study the convergence of slow incentive updates, assuming that the fast strategy updates have converged to the equilibrium.

Formally, to study the convergence of fast strategy updates, we re-write (x-update)-(p-update) as follows

$$\begin{aligned} x_{k+1} &= x_k + \gamma_k \left(f(x_k, p_k) - x_k\right) \\ p_{k+1} &= p_k + \gamma_k \frac{\beta_k}{\gamma_k}\left(e(x_k) - p_k\right). \end{aligned} \tag{2.20}$$

Since $\sup_{k\in\mathbb{N}}(\|x_k\| + \|p_k\|) < +\infty$ and $\lim_{k\to\infty}\beta_k/\gamma_k = 0$ (compare Assumption 2.4.1), the term $\frac{\beta_k}{\gamma_k}\left(e(x_k) - p_k\right)$ in (2.20) goes to zero as $k \to \infty$. Consequently, leveraging the standard approximation arguments [31, Lemma 1, Section 2.2], we conclude that the asymptotic behavior of the updates in (2.20) is the same as that of the following dynamical system:

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{p}(t)) - \mathbf{x}(t), \quad \dot{\mathbf{p}}(t) = 0.$$

Using Assumption 2.4.2, we conclude that

$$\lim_{k\to\infty}(x_k, p_k) \to \{(x^*(p), p) : p \in \mathbb{R}^{|\mathcal{I}|}\}. \tag{2.21}$$

Next, to study the convergence of the slow incentive updates, we re-write (p-update) as follows

$$p_{k+1} = p_k + \beta_k\left(e(x^*(p_k)) - p_k\right) + \beta_k\left(e(x_k) - e(x^*(p_k))\right). \tag{2.22}$$

We will show that $(p_k)_{k\in\mathbb{N}}$ will asymptotically follow the trajectories of the following continuous-time dynamical system:

$$\dot{\mathbf{p}}(t) = e(x^*(\mathbf{p}(t))) - \mathbf{p}(t). \tag{2.23}$$

Note that $p^\dagger$ is the fixed point of the trajectories of dynamical system (2.23) (compare Proposition 2). Conditions (C1) and (C2) in Proposition 3 are based on results from non-linear

[1] The results in this chapter hold even if the updates (x-update)-(p-update) and ($\tilde{x}$-update)-($\tilde{p}$-update) are perturbed with martingale difference noise [32].

dynamical systems which ensure convergence of (2.23). In particular, (C1)-(i) (respectively (C1)-(ii)) builds on cooperative dynamical systems theory [93], which ensures that $\mathbb{R}_+^{|\mathcal{I}|}$ (respectively $\mathbb{R}_-^{|\mathcal{I}|}$) is positively invariant for (2.23) and $p^\dagger \in \mathbb{R}_+^{|\mathcal{I}|}$ (respectively $p^\dagger \in \mathbb{R}_-^{|\mathcal{I}|}$) is asymptotically stable. On the other hand, condition (C2) ensures the existence of a Lyapunov function that is strictly positive everywhere except at $p^\dagger$ and decreases along the flow of (2.23) ([171]). Either condition guarantees the convergence of the trajectories of (2.23) to $p^\dagger$.

Let $D^\dagger$ denote the domain of attraction of $p^\dagger$ for the dynamical system (2.23). From the converse Lyapunov theorem [173], we know that there exists a continuously differentiable function $\bar{V} : D^\dagger \to \mathbb{R}_+$ such that $\bar{V}(p^\dagger) = 0$, $\bar{V}(p) > 0$ for all $p \in D^\dagger \backslash \{p^\dagger\}$ and $\bar{V}(p) \to \infty$ as $p \to \mathsf{boundary}(D^\dagger)$. For any $r > 0$, define $\bar{V}_r = \{p \in \mathsf{dom}(\bar{V}) : \bar{V}(p) \leq r\}$ to be a sub-level set of $\bar{V}$. There exists $0 < \bar{r}' < \bar{r}$ such that $\bar{V}_{\bar{r}'} \subsetneq \mathcal{B}_{\bar{r}'}(p^\dagger) \subsetneq \mathcal{B}_{\bar{r}}(p^\dagger) \subsetneq \bar{V}_{\bar{r}}$. Additionally, define $\ell_0 = 0, \ell_k = \sum_{i=1}^k \beta_i$ and $L_k = \ell_{n(k)}$ where $n(0) = 0$, and

$$n(k) = \min \left\{ m \geq n(k-1) : \sum_{j=n(k-1)+1}^{m} \beta_j \geq T \right\} \ \forall k \in \mathbb{N}. \tag{2.24}$$

Here, T is a positive integer to be described shortly. Furthermore, define $\bar{\mathbf{p}}^{(k)} : \mathbb{R}_+ \to \mathbb{R}^{|\mathcal{I}|}$ to be a solution of (2.23) on $[L_k, \infty)$ such that $\bar{\mathbf{p}}^{(k)}(L_k) = p_{L_k}$.

To ensure that $\bar{\mathbf{p}}^{(k)}(L_k) \in \mathsf{dom}(\bar{V})$ for $k > 0$, we show that for an appropriate choice of T in (2.24), $p_{L_k} \in \mathsf{int}(D^\dagger)$ for every $k \in \mathbb{N}$. From [32, Theorem IV.1], we know that there exists $K > 0$ such that for all $k \in \mathbb{N}$,

$$\begin{aligned} &\|p_k - \bar{\mathbf{p}}^{(0)}(\ell_k)\| \\ &\leq K \left(\sup_k \beta_k + \sup_k \gamma_k + \sup_k \frac{\beta_k}{\gamma_k} + \sup_k \frac{\beta_k}{\gamma_k} \|x_0 - x^*(p_0)\| \right) \\ &= K \left(\bar{\alpha} + \bar{\beta} + \bar{\gamma} + \bar{\alpha}\bar{r} \right) =: \kappa. \end{aligned}$$

Consequently, using the triangle inequality, it holds that

$$\|p_k - p^\dagger\| \leq \kappa + \|\bar{\mathbf{p}}^{(0)}(\ell_k) - p^\dagger\|. \tag{2.25}$$

Since $\bar{V}$ is a Lyapunov function of (2.23) and $\bar{\mathbf{p}}^{(0)}(0) = p_0 \in \mathcal{B}_{\bar{r}}(p^\dagger) \subsetneq \bar{V}_{\bar{r}}$, there exists $\bar{k} \in \mathbb{N}$ such that for all $k \geq \bar{k}$, $\bar{\mathbf{p}}^{(0)}(\ell_k) \in \bar{V}_{\bar{r}'} \subsetneq \mathcal{B}_{\bar{r}'}(p^\dagger)$. If we choose $\kappa < \bar{r} - \bar{r}'$ then, from (2.25), it holds that for all $k \geq \bar{k}$, $p_k \in \mathcal{B}_{\bar{r}}(p^\dagger)$. Therefore, if we choose $T \geq \bar{k}$ in (2.24), it holds that

$$p_{L_k} \in \mathsf{dom}(\bar{V}), \quad \forall\, k \in \mathbb{N}. \tag{2.26}$$

Define $\hat{p} : \mathbb{R}_+ \to \mathbb{R}$ such that, for every $k \in \mathbb{N}$, $\hat{p}(\ell_k) = p_k$ with linear interpolation on $[\ell_k, \ell_{k+1}]$. Using the standard approximation arguments from [31, Chapter 6], it holds that[2]

$$\sup_{t \in [L_k, L_{k+1}]} \|\hat{p}(t) - \bar{\mathbf{p}}^{(k)}(t)\|$$

[2]For any $T \geq \bar{k}$ and $\delta > 0$, there exists $k(\delta)$ such that $\hat{p}(\ell_{k(\delta)} + \cdot)$ form a "(T, δ)" perturbation (compare [30]) of (2.23).

$$\leq \mathcal{O}\left(\sum_{m\geq L_k}\beta_m^2 + \sup_{m\geq L_k}\|x_m - x^*(p_m)\|\right). \tag{2.27}$$

Using (2.21) and Assumption 2.4.1, we conclude that the right hand side in the above equation goes to zero as $k \to \infty$. Finally, using (2.26), (2.27) and [30, Lemma 2.1] we conclude that $p_k \to p^\dagger$ as $k \to \infty$.

□

2.5 Applications

In this section, we study the applicability of the proposed dynamic incentive design approach in two practically relevant classes of games: atomic aggregative games and non-atomic routing games.

Atomic Aggregative Games

First, we study *quadratic* networked aggregative games [34, 33, 174, 1]. Consider a game G comprised of a finite set of players $\mathcal{I}$. The strategy set of every player is the entire real line $\mathbb{R}$ (note that although this strategy set is not compact, compactness is only required to ensure the existence of the Nash equilirium of the underlying game; in aggregative games, we show that the Nash equilibrium exists and is unique for every incentive). Given the joint strategy profile $x = (x_i)_{i\in\mathcal{I}}$, the cost of each player $i \in \mathcal{I}$ is

$$\ell_i(x) = \frac{1}{2}q_i x_i^2 + \alpha x_i (Ax)_i, \tag{2.28}$$

where $A \in \mathbb{R}^{|\mathcal{I}|\times|\mathcal{I}|}$ is the *network matrix*, where A_{ij} represents the impact of player j's strategy on the cost of player i, and $\alpha > 0$ is a constant that characterizes the impact of the aggregate strategy on the individual cost of players. Moreover, $q_i > 0$ characterizes the impact of each player's own strategy on their cost function. Without loss of generality, we consider $A_{ii} = 0$ for all $i \in \mathcal{I}$. For notational brevity, we define $Q = \mathsf{diag}((q_i)_{i\in\mathcal{I}}) \in \mathbb{R}^{|\mathcal{I}|\times|\mathcal{I}|}$.

A system operator designs incentives by setting a payment $p_i x_i$ for player i for choosing strategy x_i. Thus, given a incentive vector $p = (p_i)_{i\in\mathcal{I}}$, the total cost of player i is given by $c_i(x,p) = \ell_i(x) + p_i x_i$. The system operator's objective is to ensure that the players reach a strategy $x^\dagger$, which minimizes the following social cost function:

$$\Phi(x) = \sum_{i=1}^{n} h_i(x_i), \tag{2.29}$$

Assumption 2.5.1. *We assume that for every $i \in \mathcal{I}$, $h_i(\cdot)$ is a strictly convex function with Lipschitz continuous gradient. Furthermore, we assume the existence of $y^\dagger \in \mathbb{R}^{|\mathcal{I}|}$ such that $\nabla h_i(y_i^\dagger) = 0$ for every $i \in \mathcal{I}$.*

Such separable social costs have been considered in the literature on control of agents in aggregative games [84] as well as systemic risk analysis in financial networks [2]. One example of $\Phi(x)$ that satisfies this property is $\Phi(x) = \frac{1}{2}\sum_{i\in\mathcal{I}}(x_i - \xi_i)^2$, which computes the difference between the players' strategies and some desired strategy $\xi \in \mathbb{R}^{|\mathcal{I}|}$.

Proposition 4. *Suppose that Assumption 2.5.1 holds and $M := Q + \alpha A$ is invertible. Then, the Nash equilibrium $x^*(p)$ exists and is unique for every incentive vector $p \in \mathbb{R}^{|\mathcal{I}|}$. Furthermore, the set $P^\dagger$ is a singleton set.*

Proof. First, we show that if M is invertible, then $x^*(p) = -M^{-1}p$ for any $p \in \mathbb{R}^{|\mathcal{I}|}$. To see this, observe that the cost function $c_i(x_i, x_{-i}, p)$ is strongly convex in x_i and the strategy space X_i is unconstrained. Therefore, it must hold that $\nabla_{x_i} c_i(x^*(p), p) = 0$, for every $i \in \mathcal{I}$. Consequently, using (2.28), it must hold that

$$q_i x_i^*(p) + \alpha(Ax^*(p))_i + p_i = 0, \quad \forall\ i \in \mathcal{I}. \tag{2.30}$$

Stacking (2.30) in a vector form yields $Mx^*(p) = -p$.

Next, we show that $P^\dagger$ is a singleton set. From (2.7a), the externality in this setup is expressed as:

$$e_i(x) = \nabla h_i(x_i) - q_i x_i - \alpha \sum_{j=1}^{n} A_{ij} x_j,\ \forall i \in \mathcal{I},\ \forall x. \tag{2.31}$$

Combining (2.30) and (2.31), we obtain

$$e_i(x^*(p)) = \nabla h_i(x_i^*(p)) + p_i. \tag{2.32}$$

Consequently, using (2.9), we obtain

$$P^\dagger = \{p^\dagger \in \mathbb{R}^{|\mathcal{I}|} : \nabla h_i(x_i^*(p^\dagger)) = 0, \ \ \forall\ i \in \mathcal{I}\}. \tag{2.33}$$

Since h_i is strictly convex, from Assumption 2.5.1, there exists a *unique* $y^\dagger$ such that $\nabla h_i(y_i^\dagger) = 0$, for every $i \in I$. Therefore, for every $p^\dagger \in P^\dagger$ it should hold that $x^*(p^\dagger) = y^\dagger$. Using the fact that $x^*(p) = -M^{-1}p$, we obtain the unique $p^\dagger = -My^\dagger$. □

Next, we provide sufficient conditions to ensure the convergence of (x-update)-(p-update).

Proposition 5. *Consider the aggregative game G with invertible M. Suppose that Assumptions 2.4.1, 2.4.2, and 2.5.1 are satisfied, and at least one of the following conditions hold:*

(A1) The matrix M has non-negative entries, M^{-1} has strictly-negative off-diagonal entries. Furthermore, there exists[3] some $y^\dagger \in \mathbb{R}^{|\mathcal{I}|}$ such that $\nabla h_i(y_i^\dagger) = 0$ for every $i \in \mathcal{I}$, and all entries of $y^\dagger$ have the same sign;

[3]Since h_i is strictly convex then $\nabla h_i(\cdot)$ is strictly increasing and can have only one zero-crossing, which ensures that $y^\dagger$ is unique.

(A2) M is a symmetric positive definite matrix.
Additionally, $\sup_{k\in\mathbb{N}}(\|x_k\| + \|p_k\|) < +\infty$*. Then there exist positive scalars $\bar{r}, \bar{\alpha}, \bar{\beta}, \bar{\gamma}$ such that for every $p_0 \in \mathcal{B}_{\bar{r}}(p^\dagger)$, $x_0 \in \mathcal{B}_{\bar{r}}(x^*(p_0))$, with step-sizes that follow* (2.18) *must satisfy* $\lim_{k\to\infty}(x_k, p_k) = (x^\dagger, p^\dagger)$.

Proof. The proof follows by verifying that the conditions $(A1)$ and $(A2)$ ensure that the conditions $(C1)$ and $(C2)$ in Proposition 3 are satisfied, respectively.

To begin, we show that condition $(A1)$ ensures that condition $(C1)$ is satisfied in Proposition 3. Recall, from (2.32), that $e_i(x^*(p)) = \nabla h_i(x_i^*(p)) + p_i$. First, we show that $i, j \in \mathcal{I}$ such that $i \neq j$, $\frac{\partial e_i(x^*(p))}{\partial p_j} > 0$. Indeed,

$$\begin{aligned}\frac{\partial e_i(x^*(p))}{\partial p_j} &= \nabla^2 h_i(x_i^*(p))\frac{\partial x_i^*(p)}{\partial p_j} \\ &= \nabla^2 h_i(x_i^*(p))(-M^{-1})_{ij} > 0,\end{aligned}$$

where the inequality holds because h_i is strictly convex and $(M^{-1})_{ij} < 0$.

Next, we show that if $e_i(x^*(0)) \geq 0$ for every $i \in \mathcal{I}$ then there exists $p \in \mathbb{R}_+^{|\mathcal{I}|}$ such that $e_i(x^*(p)) - p_i \leq 0$ for every $i \in \mathcal{I}$ (compare (C1)-(i) in Proposition 3). Using (2.32), it is sufficient to show that

$$\begin{aligned}\nabla h_i(0) \geq 0 \quad \forall\, i \in \mathcal{I} \implies &\exists\, \bar{p} \in \mathbb{R}_+^{|\mathcal{I}|} \text{s.t.} \\ &\nabla h_i(x^*(\bar{p})) \leq 0 \quad \forall\, i \in \mathcal{I}.\end{aligned} \tag{2.34}$$

By $(A1)$, we know that $\nabla h_i(y_i^\dagger) = 0$ and $y_i^\dagger y_j^\dagger \geq 0$ for every $i, j \in \mathcal{I}$. We claim that $y_i^\dagger \nabla h_i(0) \leq 0$ for every $i \in \mathcal{I}$. Indeed, for every $i \in \mathcal{I}$,

$$0 \leq (\nabla h_i(y_i^\dagger) - \nabla h_i(x_i^*(0)))(y_i^\dagger - x_i^*(0)) = -y_i^\dagger \nabla h_i(0),$$

where the first inequality is due to the strict convexity of h_i. Suppose $\nabla h_i(0) \geq 0$, for every $i \in \mathcal{I}$, then $-y^\dagger \in \mathbb{R}_+^{|\mathcal{I}|}$. Next, define $\bar{p} = -(1+\epsilon)My^\dagger$, for some $\epsilon > 0$, then $\bar{p} \in \mathbb{R}_+^{|\mathcal{I}|}$ as $y^\dagger \in \mathbb{R}_+^{|\mathcal{I}|}$ and M has all non-negative entries (compare $(A1)$). This ensures that $x^*(\bar{p}) = (1+\epsilon)y^\dagger$. We claim that $\nabla h_i(x_i^*(\bar{p})) < 0$ for every $i \in \mathcal{I}$. Indeed, for every $i \in \mathcal{I}$,

$$\begin{aligned}0 &< (\nabla h_i(x_i^*(\bar{p})) - \nabla h_i(y_i^\dagger))(x_i^*(\bar{p}) - y_i^\dagger) \\ &= \nabla h_i(x_i^*(\bar{p}))\epsilon y_i^\dagger.\end{aligned}$$

The claim follows because $y_i^\dagger \leq 0$ and $\epsilon > 0$. This shows that the implication in (2.34) holds. Similarly, it can be show that if $e_i(x^*(0)) \leq 0$ for every $i \in \mathcal{I}$ then there exists $p \in \mathbb{R}_-^{|\mathcal{I}|}$ such that $e_i(x^*(p)) - p_i \geq 0$ for every $i \in \mathcal{I}$ (compare (C1)-(ii) in Proposition 3).

Next, we verify that condition $(A2)$ ensures that condition $(C2)$ in Proposition 3 is satisfied. Consider the function $V(p) = (p - p^\dagger)^\top M^{-\top}(p - p^\dagger)$. Note that $V(p^\dagger) = 0$ and $V(p) > 0$ for all $p \neq p^\dagger$. Next, we show that $\nabla V(p)^\top(e(x^*(p)) - p) \leq 0$. Indeed,

$$
\begin{aligned}
&\nabla V(p)^\top(e(x^*(p)) - p) = 2(p - p^\dagger)^\top M^{-\top}\nabla h(x^*(p)) \\
&= -2(x^*(p) - x^*(p^\dagger))\nabla h(x^*(p)) \\
&= -2(x^*(p) - x^*(p^\dagger))\left(\nabla h(x^*(p)) - \nabla h(x^*(p^\dagger))\right) \\
&= -2(x^*(p) - x^*(p^\dagger))\left(\nabla h(x^*(p)) - \nabla h(x^*(p^\dagger))\right) \\
&\leq 0
\end{aligned}
$$

where the last equality holds due to the strict convexity of h_i for every $i \in \mathcal{I}$. □

Remark 2.5.1. *There exist instances of quadratic networked aggregative games such that only one of the conditions* $(A1)$ *or* $(A2)$ *is satisfied in Proposition 5. For instance consider the matrices*

$$M_1 = \begin{bmatrix} 1 & 0.1 \\ 1 & 1 \end{bmatrix}, M_2 = \begin{bmatrix} 1 & -0.1 \\ -0.1 & 1 \end{bmatrix}.$$

Note that M_1 *satisfies* $(A1)$ *but not* $(A2)$ *and the matrix* M_2 *satisfies* $(A2)$ *but not* $(A1)$*.*

Non-atomic Traffic Routing on General Networks

Consider a routing game $\tilde{G}$ characterizing the interaction of strategic travelers over a graph $\tilde{\mathcal{G}} = (\tilde{\mathcal{E}}, \tilde{\mathcal{N}})$ where $\tilde{\mathcal{N}}$ is the set of nodes and $\tilde{\mathcal{E}}$ denotes the set of edges. Let $\tilde{\mathcal{I}}$ be the set of origin-destination (o-d) pairs. Each o-d pair $i \in \tilde{\mathcal{I}}$ is connected by a set of routes[4] represented by $\tilde{R}_i$. Let $\tilde{R} = \cup_{i\in\tilde{\mathcal{I}}}\tilde{R}_i$ denote the set of all routes on the network.

An infinitesimal traveler on the network is associated with an o-d pair and chooses a route to commute between the o-d pair. Let the total population of travelers associated with any o-d pair $i \in \tilde{\mathcal{I}}$ be denote by $\tilde{M}_i$. Let $\tilde{x}_i^j$ be the amount of travelers taking route $j \in \tilde{R}_i$ to commute between o-d pair $i \in \tilde{\mathcal{I}}$ and $\tilde{x} = (\tilde{x}_i^j)_{j\in\tilde{R}_i, i\in\tilde{\mathcal{I}}}$ is a vector which contains, as its entries, the route flow of all population on different routes. Naturally, it holds that $\sum_{j\in\tilde{R}_i}\tilde{x}_i^j = \tilde{M}_i$ for every $i \in \tilde{\mathcal{I}}$. Any route flow $\tilde{x}$ induces a flow on the edges of the network denoted by $\tilde{w}$ such that $\tilde{w}_a = \sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}\tilde{x}_i^j\mathbb{1}(a \in j)$ for every $a \in \tilde{\mathcal{E}}$. We denote the set of feasible route flows by $\tilde{X}$ and the set of feasible edge flows by $\tilde{W} = \{(\tilde{w}_a)_{a\in\tilde{\mathcal{E}}} : \exists\tilde{x} \in \tilde{X}, \tilde{w}_a = \sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}\tilde{x}_i^j\}$.

For any o-d pair $i \in \tilde{\mathcal{I}}$ and route flow $\tilde{x} \in \mathbb{R}^{|\tilde{R}|}$, the cost experienced by travelers using route $j \in \tilde{R}_i$ is $\tilde{\ell}_i^j(\tilde{x}) = \sum_{a\in\tilde{\mathcal{E}}} l_a(\tilde{w}_a)\mathbb{1}(a \in j)$, where $l_a(\cdot)$ is the *edge latency function* that depends on the edge flows. For every edge $a \in \tilde{\mathcal{E}}$, we assume that the edge latency function $l_a(\cdot)$ is convex and strictly increasing. This property of edge latency function captures the congestion effect on the transportation network [22, 169].

[4] A route is a set of contiguous edges.

A system operator designs incentives by setting tolls on the edges of the network in the form of edge tolls[5], denoted by $\tilde{p} = (\tilde{p}_a)_{a\in\tilde{\mathcal{E}}}$. Every edge toll vector induces a unique route toll vector $\tilde{P}$. That is, for any o-d pair $i \in \tilde{\mathcal{I}}$ the toll on route $j \in \tilde{R}_i$ is

$$\tilde{P}_i^j = \sum_{a\in\tilde{\mathcal{E}}:a\in j} \tilde{p}_a. \tag{2.35}$$

Consequently, the total cost experienced by travelers on o-d pair $i \in \tilde{\mathcal{I}}$ who choose route $j \in \tilde{R}_i$ is $\tilde{c}_i^j(\tilde{x}, \tilde{P}) = \tilde{\ell}_i^j(\tilde{x}) + \tilde{P}_i^j$. Let $\tilde{x}^*(\tilde{P})$ denote a Nash equilibrium corresponding to route tolls $\tilde{P}$. Owing to (2.35), with slight abuse of notation, we shall frequently use $\tilde{x}^*(\tilde{P})$ and $\tilde{x}^*(\tilde{p})$ interchangeably. Typically, the equilibrium route flows can be non-unique but the corresponding edge flows $\tilde{w}^*(\tilde{p})$ are unique. Furthermore, the function $\tilde{p} \mapsto \tilde{w}^*(\tilde{p})$ is a continuous function [191].

The system operator's objective is to design tolls that ensure that the resulting equilibrium minimizes the overall travel time incurred by travelers on the network, characterized as the minimizer of

$$\tilde{\Phi}(\tilde{x}) = \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \tilde{x}_i^j \tilde{\ell}_i^j(\tilde{x}). \tag{2.36}$$

Note that the optimal route flow can be non-unique but the optimal edge flow, denoted by $w^\dagger$, is unique [191].

Using the description of cost of travelers, the externality, based on (2.7b), caused by travelers from o-d pair $i \in \tilde{\mathcal{I}}$ using the route $j \in \tilde{R}_i$ is

$$\tilde{e}_i^j(\tilde{x}) = \sum_{i'\in\tilde{\mathcal{I}}} \sum_{j'\in\tilde{R}_i} \tilde{x}_{i'}^{j'} \frac{\partial \tilde{\ell}_{i'}^{j'}(\tilde{x})}{\partial \tilde{x}_i^j} \overset{(a)}{=} \sum_{a\in\tilde{\mathcal{E}}:a\in j} \nabla l_a(\tilde{w}_a)\tilde{w}_a, \tag{2.37}$$

where (a) is due to Lemma A.0.2 in Appendix A.

Define

$$\tilde{\mathbf{P}}^\dagger = \{(\tilde{p}_a^\dagger)_{a\in\tilde{\mathcal{E}}} : \tilde{p}_a^\dagger = \tilde{w}_a^*(\tilde{p}^\dagger)\nabla l_a(\tilde{w}_a^*(\tilde{p}^\dagger)), \forall\, a \in \tilde{\mathcal{E}}\}.$$

Proposition 6. *The set $\tilde{\boldsymbol{P}}^\dagger$ is a non-empty singleton set. The unique $p^\dagger \in \tilde{\boldsymbol{P}}^\dagger$ is socially optimal, i.e. $\tilde{w}(p^\dagger) = w^\dagger$.*

Proof. First, we show that $\tilde{\mathbf{P}}^\dagger$ is non-empty. This can be shown analogously to the proof of existence in Proposition 2 by using the Schauder fixed point theorem and the continuity of the function $\tilde{w}^*(\cdot)$. We omit the details of this proof.

[5] If we directly use the setup of non-atomic games presented in Section 2.3 we would require the system operator to use *route-based* tolls rather than *edge-based* tolls. This is often impractical, as the set of routes can be very large.

Next, we show that any $p^\dagger \in \tilde{\mathbf{P}}^\dagger$ aligns the Nash equilibrium with social optimality, i.e. $\tilde{w}(p^\dagger) = w^\dagger$. For any $p^\dagger \in \tilde{\mathbf{P}}^\dagger$, we have $\tilde{p}_a^\dagger = \tilde{w}_a^*(\tilde{p}^\dagger)\nabla l_a(\tilde{w}_a^*(\tilde{p}^\dagger))$ for every $a \in \tilde{\mathcal{E}}$. This implies,

$$\frac{\partial}{\partial \tilde{w}_a}(\tilde{w}_a(\tilde{p}^\dagger)l_a(\tilde{w}_a(\tilde{p}^\dagger))) = l_a(\tilde{w}_a(\tilde{p}^\dagger)) + \tilde{p}_a^\dagger, \quad \forall\ a \in \tilde{\mathcal{E}}. \tag{2.38}$$

Note that for any arbitrary edge toll $\tilde{p} \in \mathbb{R}^{|\tilde{\mathcal{E}}|}$, $\tilde{w}^*(\tilde{p})$ is the unique solution to the following strictly convex optimization problem [170].

$$\min_{\tilde{w} \in \tilde{W}} \quad \tilde{T}(\tilde{w}) = \sum_{a \in \tilde{\mathcal{E}}} \int_0^{\tilde{w}_a} l_a(\tau) d\tau + \sum_{a \in \tilde{\mathcal{E}}} \tilde{p}_a \tilde{w}_a. \tag{2.39}$$

Therefore, $\tilde{w}^*(\tilde{p})$ is a Nash equilibrium if and only if

$$\sum_{a \in \tilde{\mathcal{E}}} (l_a(\tilde{w}_a(\tilde{p}) + \tilde{p}_a)(\tilde{w}_a - \tilde{w}_a(\tilde{p}))) \geq 0 \quad \forall\ \tilde{w} \in \tilde{W}. \tag{2.40}$$

Combining (2.38) and (2.40), we conclude that

$$\sum_{a \in \tilde{\mathcal{E}}} \frac{\partial}{\partial \tilde{w}_a}(\tilde{w}_a(\tilde{p}^\dagger)l_a(\tilde{w}_a(\tilde{p}^\dagger)))(\tilde{w}_a - \tilde{w}_a^*(\tilde{p}^\dagger)) \geq 0 \quad \forall\ \tilde{w} \in \tilde{W}. \tag{2.41}$$

Further, from the first order conditions of optimality for the social cost function, we know that $\tilde{x}^\dagger$ is socially optimal if and only if

$$\sum_{i \in \tilde{\mathcal{I}}} \sum_{j \in \tilde{R}_i} \frac{\partial \Phi(\tilde{x}^\dagger)}{\partial \tilde{x}_i^j}(\tilde{x}_i^j - \tilde{x}_i^{\dagger j}) \geq 0. \tag{2.42}$$

Using Lemma A.0.3 in Appendix A, we can equivalently write (2.42) in terms of edge flows as follows

$$\sum_{a \in \tilde{\mathcal{E}}} \frac{\partial}{\partial \tilde{w}_a}(\tilde{w}_a^\dagger l_a(\tilde{w}_a^\dagger))(\tilde{w}_a - \tilde{w}_a^\dagger) \geq 0 \quad \forall\ \tilde{w} \in \tilde{W}, \tag{2.43}$$

where $w^\dagger$ is the edge flow corresponding to the route flow $x^\dagger$. Comparing (2.41) with (2.43), we note that $\tilde{w}^*(p^\dagger)$ is the minimizer of social cost function $\tilde{\Phi}$. This implies $\tilde{w}^*(\tilde{p}^\dagger) = \tilde{w}^\dagger$.

The proof that $\tilde{\mathbf{P}}^\dagger$ is singleton follows by contradiction which is analogous to that of Proposition 2. We omit the details.

□

Next, we show that the updates ($\tilde{x}$-update)-($\tilde{p}$-update) converge to the fixed points discussed in previous section.

Proposition 7. *Suppose that for every* $i \in \tilde{I}, j \in \tilde{R}_i$

$$\tilde{c}_i^j(\tilde{x}^*(\tilde{p}^\dagger)) \leq \tilde{c}_i^{j'}(\tilde{x}^*(\tilde{p}^\dagger)) \;\forall\; j' \in \tilde{R}_i \implies \tilde{x}^*_{i,j}(\tilde{p}^\dagger) > 0. \tag{2.44}$$

Additionally, $\sup_{k\in\mathbb{N}}(\|\tilde{x}_k\| + \|\tilde{p}_k\|) < +\infty$*. Then there exist positive scalars* $\bar{r}, \bar{\alpha}, \bar{\beta}, \bar{\gamma}$ *such that for every* $\tilde{p}_0 \in \mathcal{B}_{\bar{r}}(\tilde{p}^\dagger)$*,* $\tilde{x}_0 \in \mathcal{B}_{\bar{r}}(\tilde{x}^*(\tilde{p}_0))$*, with step-sizes that follow* (2.18) *must satisfy* $\lim_{k\to\infty}(\tilde{x}_k, \tilde{p}_k) = (\tilde{x}^\dagger, \tilde{p}^\dagger)$*.*

Proof. The proof follows by verifying the second convergence condition in Proposition 3. We show that $V(\tilde{p}) = (\tilde{p} - \tilde{p}^\dagger)^\top \Delta(\tilde{p} - \tilde{p}^\dagger)$ serves as candidate Lyapunov function which satisfies the conditions stated in Proposition 3-(C2) if we set $\Delta \in \mathbb{R}^{|\tilde{\mathcal{E}}|\times|\tilde{\mathcal{E}}|}$ to be a diagonal matrix such that

$$\Delta_{a,a} = (\nabla l_a(\tilde{w}_a^*(\tilde{p}^\dagger)) + \tilde{w}_a^*(\tilde{p}^\dagger)\nabla^2 l_a(\tilde{w}_a^*(\tilde{p}^\dagger)))^{-1}.$$

Due to the strict monotonicity and convexity of $l_a(\cdot)$, it follows that $\Delta_{a,a} > 0$ for every $a \in \tilde{\mathcal{E}}$. Towards this goal, we show that there exists a positive scalar r such that for any $\tilde{p} \in \mathcal{B}_r(\tilde{p}^\dagger)$, it holds that

$$\sum_{a\in\tilde{\mathcal{E}}} \nabla_{\tilde{p}_a} V(\tilde{p})^\top \left(\tilde{w}_a^*(\tilde{p})\nabla l_a(\tilde{w}_a^*(\tilde{p})) - \tilde{p}_a\right) < -2V(\tilde{p}). \tag{2.45}$$

Indeed, we note that

$$\begin{aligned}
&\sum_{a\in\tilde{\mathcal{E}}} \nabla_{\tilde{p}_a} V(\tilde{p}) \left(\tilde{w}_a^*(\tilde{p})\nabla l_a(\tilde{w}_a^*(\tilde{p})) - \tilde{p}_a\right) \\
&= 2\sum_{a\in\tilde{\mathcal{E}}} \Delta_{a,a}(\tilde{p}_a - \tilde{p}_a^\dagger) \left(\tilde{w}_a^*(\tilde{p})\nabla l_a(\tilde{w}_a^*(\tilde{p})) - \tilde{p}_a\right) \\
&= 2\sum_{a\in\tilde{\mathcal{E}}} \Delta_{a,a}(\tilde{p}_a - \tilde{p}_a^\dagger) \left(\tilde{w}_a^*(\tilde{p})\nabla l_a(\tilde{w}_a^*(\tilde{p})) - \tilde{p}_a^\dagger + \tilde{p}_a^\dagger - \tilde{p}_a\right) \\
&= -2V(\tilde{p}) + 2\sum_{a\in\tilde{\mathcal{E}}} \Delta_{a,a}(\tilde{p}_a - \tilde{p}_a^\dagger) \left(\phi_a(\tilde{p}) - \phi_a(\tilde{p}^\dagger)\right)
\end{aligned}$$

where for every $a \in \tilde{\mathcal{E}}$, $\phi_a(\tilde{p}) := \tilde{w}_a^*(\tilde{p})\nabla l_a(\tilde{w}_a^*(\tilde{p}))$. Thus it suffices to show that

$$\sum_{a\in\tilde{\mathcal{E}}} \Delta_{a,a}(\tilde{p}_a - \tilde{p}_a^\dagger) \left(\phi_a(\tilde{p}) - \phi_a(\tilde{p}^\dagger)\right) \leq 0. \tag{2.46}$$

Condition (2.44) ensures that the function $\tilde{\phi}$ is differentiable in a neighborhood of $\tilde{p}^\dagger$ (compare [191, Chapter 4]). From Lemma A.0.1 in Appendix A, we know that if $z^\top \Delta D\phi(\tilde{p}^\dagger)z \leq 0$ for all z then there exists $r > 0$ such that (2.46) holds for all $\tilde{p} \in \mathcal{B}_r(\tilde{p}^\dagger)$. Indeed, by the design of Δ it holds that $\Delta D\phi(\tilde{p}^\dagger) = D\tilde{w}(\tilde{p}^\dagger)$. Finally, using Lemma A.0.1 and Lemma A.0.4 in Appendix A, we note that $z^\top D\tilde{w}(\tilde{p}^\dagger)z \leq 0$. This concludes the proof. □

2.6 Discussion and Future Work

We propose an adaptive incentive mechanism that (*i*) updates based on the feedback of each agent's externality, (*ii*) is agnostic to the learning rule employed by agents, and (*iii*) updates at a slower timescale compared to the agents' learning dynamics, leading to a two-timescale coupled dynamical system. We provide sufficient conditions that ensure the convergence of this dynamical system to the optimal incentive mechanism, guaranteeing that the Nash equilibrium of the corresponding game among agents achieves social optimality. We demonstrate the effectiveness of our proposed incentive mechanism in two practically relevant games: atomic networked quadratic aggregative games and non-atomic network routing games, by verifying the proposed convergence conditions.

An important direction for future research is to verify the convergence conditions in a broader class of games. Notably, leveraging the externality-based dynamic incentive design update proposed in an earlier version of this chapter (compare [133]), [122] studied the convergence of such updates in power systems.

Chapter 3

Causal Discovery for Rare Events

It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.

Mark Twain

This chapter can be found in [40].

3.1 Introduction

The occurrence of rare yet consequential events during the evolution of a dynamical system is ubiquitous in many fields of engineering and science. Examples include natural disasters, vehicular accidents, and stock market crashes. When studying such phenomena, it is crucial to understand the causal links between the disruptive event and the underlying system dynamics. In particular, if certain values of the system state increase the probability that the disruptive event occurs, control strategies should be implemented to steer the state away from such values. This can be accomplished, for instance, by incorporating a description of this causal relationship into the cost function that generates these control inputs in an optimization-based control framework. In general, it is important to consider the following question:

Main Question (Q): Given a rare event associated with a dynamical system, does the onset of the event become more likely when the system state assumes certain values?

Below, we present a running example, invoked throughout ensuing sections to provide context.

Running Example: Consider the task of reducing the number of vehicular accidents on a road by identifying their causes. In particular, consider the scenario in which the amount of traffic on a network of roads has a causal effect on accident occurrence. For example, on busy streets, high traffic flow may render chain collisions more likely. In this case, since steady-state flows in a traffic network can be controlled via incentives, regulators can adjust prices on each network link to redistribute flow and reduce the number of accidents that transpire [130, 131]. Conversely, on other roads, low traffic flow may incentivize drivers to exceed the speed limit and create more opportunities for accidents to occur. In this case, traffic engineers can enforce speed limits more stringently at times of low traffic flow.

Although many well-established methods in the causal discovery literature can efficiently learn causal relationships from data, most only apply to data generated from probability distributions associated with static, acyclic Bayesian networks [79, 157]. Moreover, most causal discovery algorithms developed for time series data rely on stringent assumptions, such as linear dynamics and additive Gaussian noise models, or aggregate data along slices of fixed time indices [79, 80, 85, 160]. However, rare events often occur sparsely at any fixed time and cannot be easily modeled using linear dynamics.

To address these shortcomings, we present a novel approach for aggregating and analyzing time series data recording sparsely occurring, but consequential events, in which data is collected in a time-ordered fashion from a dynamical system. Our method rests on the observation that, whereas a rare event may be highly unlikely to occur *at any fixed time* t, the probability of the event occurring *at some time along the entire horizon* of interest is often much higher. Thus, we aggregate the time series data along the times of the event's first occurrence. This renders the dataset more informative, by better representing the rare events of interest. Next, we present an algorithm that uses the curated data to analyze the causal relationships governing the occurrence of the rare event. We formally pose the question of whether the system state causally affects the occurrence of the rare event as a binary hypothesis test, with the null hypothesis H_0 corresponding to the negative answer, and the alternative hypothesis H_1 corresponding to the positive one. We prove that our proposed method is *consistent against all alternatives* [118]. In other words, if H_0 were true, then as the number of data trajectories N in the dataset approaches infinity, our approach would reject H_1 with probability 1. We validate the performance of our algorithm on simulated and on publicly available traffic and incident data collected from the Caltrans Performance Measurement System (PeMS).

3.2 Related Works

Causal Discovery for Static and Time Series Data

Causal discovery algorithms identify causal links among a collection of random variables from a dataset of their realizations. Common approaches include constraint-based methods (which use statistical independence tests), score-based methods (which pose causal discovery as an optimization problem), and hybrid methods [79, 157, 161]. However, most of these approaches apply only to non-temporal settings. For time series data, Granger causality uses vector autoregression to study whether one time series can be used to predict another [85]. Other methods aggregate different data trajectories by matching time indices [160, 60], or directly solve a time-varying causal graph [135]. However, these methods do not address the problem of inferring causal links between rare events and dynamical systems, across sample trajectories on which the rare event can often occur at different times.

Extreme Value Theory and Analysis of Rare Events

Extreme value theory characterizes dependences between random variables that exist only when a low-probability event occurs, e.g., rare meteorological events, or financial crises [59, 12]. Most closely related to our work are [80], which studies causal links between heavy-tailed random variables, and [101], which explores causal relationships between characteristics of London bicycle lanes, e.g., density, length, and collision rate, and abnormal congestion. However, [80] imposes restrictive assumptions, such as linear models, while the discussion in [101] on accidents' occurrences is restricted to empirical studies. In contrast, our proposed algorithm returns a nonparametric conditional independence test statistic that is capable of characterizing relationships between a general dynamical system, and the onset of a rare event.

Traffic Network Analysis

Traffic network theory aims to mathematically describe and control traffic flow in urban networks of roads, bridges, and highways [17, 105, 4]. Recent literature has proposed the design of tolling mechanisms that drive a traffic network to the socially optimal steady state [131, 44]. However, these methods do not model or predict the occurrence of sudden yet consequential events, such as extreme weather events, car accidents, and other causes of unexpected congestion. In contrast, this chapter uses the occurrence of rare but consequential car accidents in traffic networks as a running example, to illustrate the applicability of our method on analyzing causal links between dynamical systems and associated rare events.

3.3 Preliminaries

Consider a stochastic, discrete-time dynamical system with state variable $X_t \in \mathbb{R}^n$, event variable $A_t \in \{0, 1\}$ with $\mathbb{P}(A_t = 1) \in [p_1, p_2]$ for some $p_1, p_2 \in (0, 1)$ for all t, with $p_1 < p_2$, and dynamics $X_{t+1} = f(X_t, A_t, W_t)$ for each $t \geq 0$, where $W_t \in \mathbb{R}^w$ denotes i.i.d. noise, and $f : \mathbb{R}^n \times \{0, 1\} \times \mathbb{R}^w \to \mathbb{R}^n$ denotes the nonlinear dynamics of the system state. Let T denote the time at which the rare event first occurs, and, with a slight abuse of notation, let $A_{1:t} = 0$ denote the event that $A_1 = \cdots = A_t = 0$. Moreover, we assume that the first occurrence of the rare event is governed by a time-invariant probability distribution, i.e.:

$$\mathbb{P}(A_{t+1} = 1 | X_t \preceq x, A_{1:t} = 0) \tag{3.1}$$

$$= \mathbb{P}(A_{t'+1} = 1 | X_{t'} \preceq x, A_{1:t'} = 0), \quad \forall t, t' \geq 0,$$

where, for each $x, y \in \mathbb{R}^n$, the notation $x \preceq y$ represents $x_i \leq y_i$ for each $i \in [n] := \{1, \cdots, n\}$, and for each $x \in \mathbb{R}^n$, there exists some constant ratio $\alpha(x) > 0$ such that:

$$\begin{aligned} &\mathbb{P}(X_{t-1} \preceq x | A_t = 1, A_{1:t-1} = 0) \\ =&\alpha(x) \cdot \mathbb{P}(X_{t-1} \preceq x | A_{1:t-1} = 0). \end{aligned} \tag{3.2}$$

In words, we assume that the flow distribution is related to the first occurrence of the rare event in a time-invariant manner. Given this setup, we restate **Q**, first defined in the introduction, as the following hypothesis testing problem. The binary hypothesis test, with null hypothesis H_0 as below, is a mathematically rigorous characterization of **Q**.

Definition 3.3.1. *Let H_0 be the null hypothesis given by:*

$$\begin{aligned} H_0 : \quad &\mathbb{P}(A_{t+1} = 1 | X_t \preceq x, A_{1:t} = 0) \\ =&\mathbb{P}(A_{t+1} = 1 | A_{1:t} = 0), \quad \forall x \in \mathbb{R}, \\ H_1 : \quad &\mathbb{P}(A_{t+1} = 1 | X_t \preceq x, A_{1:t} = 0) \\ \neq&\mathbb{P}(A_{t+1} = 1 | A_{1:t} = 0), \quad \forall x \in \mathbb{R}. \end{aligned}$$

In words, H_0 holds if and only if the first occurrence of the rare event transpires independently of the system state at that time. For convenience, we define the left and right hand sides of H_0 by:

$$a_1(x) := \mathbb{P}(A_{t+1} = 1 | X_t \preceq x, A_{1:t} = 0), \tag{3.3}$$

$$a_2 := \mathbb{P}(A_{t+1} = 1 | A_{1:t} = 0). \tag{3.4}$$

Running Example: Consider a parallel link traffic network of R links that connect a single source and a single destination. Let $X_{t,i} \in \mathbb{R}$ denote the traffic flow on every link $i \in [R] := \{1, \cdots, R\}$ at time t, and define $X_t := (X_{t,1}, \cdots, X_{t,r}) \in \mathbb{R}^R$. (In general, one can define $X_{t,i} \in \mathbb{R}^d$ to encapsulate other observed quantities relevant to link i at time t, e.g., vehicle speed and pavement quality). The event variable $A_t = 1$ corresponds to the occurrence of an accident in the network at time t.

In this context, Definition 3.3.1 corresponds to checking whether the first occurrence of an accident on the R-link network at time t is affected by the flow level at time $t-1$. This is of interest to traffic authorities, since costly accidents become more likely at certain levels of traffic flow X_t, then the flow should be monitored to decrease the chance that such accidents occur. Flow management can be applied by dynamically tolling the links, as in [130]. As accidents are relatively rare in most traffic datasets, it can be difficult to construct accurate estimates of accident probabilities and flows before accidents at any given time t. Instead, below, we propose a novel method of data aggregation that allows the use of information on accident occurrences across all times.

Since X_t is a continuous random variable, a direct comparison of (3.3) and (3.4) would necessitate computing (3.3) for uncountably many values of $x \in \mathbb{R}^n$. Instead, we use the laws of conditional and total probability to reformulate the problem. In the spirit of Bayes' rule, we compare the state distribution immediately before the rare event occurred, instead of the rare event probabilities under different state values. Formally, under either hypothesis, the state distribution immediately before the first accident can be decomposed as the following infinite sum; for each $x \in \mathbb{R}^n$:

$$\begin{aligned}
\mathbb{P}(X_{T-1} \preceq x) &= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, T = t) \\
&= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_t = 1, A_{1:t-1} = 0) \\
&= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \\
&\qquad \cdot \mathbb{P}(A_t = 1 | X_{t-1} \preceq x, A_{1:t-1} = 0).
\end{aligned}$$

Intuitively, if H_0 were true, then the condition $X_{t-1} \preceq x$ in the term $\mathbb{P}(A_t = 1 | X_{t-1} \preceq x, A_{1:t-1} = 0)$ can be dropped. A rigorous formulation is given in Proposition 3.3.2 below.

Proposition 3.3.2. *The null hypothesis H_0 in Definition 3.3.1 holds if and only if, for each $x \in \mathbb{R}^n$:*

$$\begin{aligned}
&\mathbb{P}(X_{T-1} \preceq x) \\
&= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \cdot \mathbb{P}(A_t = 1 | A_{1:t-1} = 0).
\end{aligned} \tag{3.5}$$

Proof. See Appendix B. □

For convenience, we define, for each $t \in \mathbb{N}$ and $x \in \mathbb{R}$:

$$\begin{aligned}
b_1(x) &:= \mathbb{P}(X_{T-1} \preceq x), \\
\beta_t(x) &:= \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0), \\
\gamma_t &:= \mathbb{P}(A_t = 1 | A_{1:t-1} = 0), \\
b_2(x) &:= \sum_{t=1}^{\infty} \beta_t(x) \cdot \gamma_t \\
&= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \\
&\qquad \cdot \mathbb{P}(A_t = 1 | A_{1:t-1} = 0)
\end{aligned}$$

$$= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \cdot \frac{\mathbb{P}(A_t = 1, A_{1:t-1} = 0)}{\mathbb{P}(A_{1:t-1} = 0)}$$
$$= \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x | A_{1:t-1} = 0) \cdot \mathbb{P}(T = t).$$

Note that $b_1(x), b_2(x) \in [0, 1]$ (in particular, that $b_2(x) \leq 1$ follows by observing that $b_2(x) \leq \sum_{t=1}^{\infty} \mathbb{P}(T = t) = 1$.)

The test statistic that we use to distinguish between the distributions $b_1(x)$ and $b_2(x)$ is the gap:

$$\sup_{x \in \mathbb{R}^n} |b_1(x) - b_2(x)|$$

Intuitively, a large gap would indicate a higher likelihood that a component-wise larger or smaller state would change the probability of an event occurring. We formalize this notion in Algorithm 1, and provide finite sample guarantees for empirical estimates of $b_1(x)$ and $b_2(x)$ that can be constructed efficiently from data and used to compute the test statistic.

Running Example: In the traffic network example, $b_1(x)$ corresponds to the probability that X_{T-1}, the network flows before the first accident, is component-wise less than or equal to x. Meanwhile, $b_2(x)$ describes the weighted average of traffic flows at each time t, conditioned on the first accident occurring after t (i.e., no accident occurs before), with the distribution of the first accident time T as weights. Section 3.4 describes sample-efficient methods for constructing empirical estimates of $b_1(x)$ and $b_2(x)$ from a dataset of independent traffic flows.

3.4 Methods

Main Algorithm

We present Algorithm 1, which solves the hypothesis testing problem in Definition 3.3.1 from a dataset of N independent trajectories, by constructing and comparing finite-sample empirical cumulative distribution functions (CDFs) $\hat{b}_1^N(x)$ and $\hat{b}_2^N(x)$ for the expressions $b_1(x)$ and $b_2(x)$, respectively, and verifying whether or not (3.5) holds (in accordance with Proposition 3.3.2).

Note on the baseline method The common baseline method for resolving the problem in Definition 3.3.1 is to fix $t \geq 1$, and compare the CDF values $\mathbb{P}(X_{t-1} \preceq x | T = t)$ and $\mathbb{P}(X_{t-1} \preceq x)$, for each $x \in \mathbb{R}^n$ at the fixed t. This is effectively a "static variant" of Algorithm 1 that only utilizes dynamical state values immediately before accidents that occur at time

t. It is generally difficult to estimate $\mathbb{P}(X_{t-1} \preceq x | T = t)$ from data, since $\mathbb{P}(T = t)$ can be very small for any given t. Our algorithm (Algorithm 1) instead aggregates data *across* times when the rare event has occurred, allowing the event to be represented with higher probability.

Algorithm 1: Hypothesis Testing with Reorganized Dataset.

Data: Dataset of system state and rare event variables: $\{(X_t^i, A_t^i) : t \geq 0, i \in [N]\}$

Result: Distribution gap: $\sup_{x \in \mathbb{R}} |\hat{b}_1^N(x) - \hat{b}_2^N(x)|$

$\hat{T}^i \leftarrow$ Realization of T for data trajectory i, $\forall i \in [N]$.

$\hat{b}_1^N(x) \leftarrow \frac{1}{N} \sum_{i=1}^N \mathbf{1}\{X_{\hat{T}^i - 1} \preceq x\}$.

$\hat{\beta}_t^N(x) \leftarrow \frac{1}{N} \sum_{i=1}^N \mathbf{1}\{X_t^i \preceq x, A_{1:t-1}^i = 0\}$.

$\hat{\gamma}_t^N \leftarrow \begin{cases} \frac{\sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0, A_t^i = 1\}}{\sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0\}}, & \text{if } \sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0\} > 0, \\ 0, & \text{else.} \end{cases}$.

$\hat{b}_2^N(x) \leftarrow \sum_{t=1}^\infty \hat{\beta}_t^N(x) \cdot \hat{\gamma}_t^N$.

return $\sup_{x \in \mathbb{R}} |\hat{b}_1^N(x) - \hat{b}_2^N(x)|$.

Theoretical Guarantees

Theorem 3.4.1 below illustrates that, if H_0 holds, then as the number of sample trajectories N approaches infinity, the empirical distributions of (B.1) and (B.2), as constructed in Algorithm 1, converge at an exponential rate to their true values. In other words, if H_0 holds, then for any fixed significance level α, Algorithm 1 will require a dataset of size no greater than $O(\ln(1/\alpha))$ to reject H_1. This establishes a finite sample bound that controls the error of the statistical independence test corresponding to the test statistic presented in Algorithm 1. The proof follows by carefully applying concentration bounds for light-tailed random variables, and invoking the Dvoretsky-Kiefer-Wolfowitz (DKW) inequality [55], which prescribes explicit convergence rates for empirical CDFs to the true CDF.

Theorem 3.4.1. ***(Exponential Convergence to Consistency Against all Alternatives)*** *Suppose the null hypothesis H_0 holds, i.e., $b_1(x) = b_2(x)$.*

1. *If $n = 1$, i.e., $X_t \in \mathbb{R}$ for each $t \geq 0$, then for each $\epsilon > 0$, there exist continuous, positive functions $C_1(\epsilon), C_2(\epsilon) > 0$ such that:*

$$\mathbb{P}\left(\sup_{x \in \mathbb{R}^n} \left\{ \left| \hat{b}_1^N(x) - \hat{b}_2^N(x) \right| \right\} > \epsilon \right) \\ \leq C_1(\epsilon) \cdot e^{-N \cdot C_2(\epsilon)}.$$

2. *If $n > 1$, then there exist continuous, positive functions $C_3(\epsilon), C_4(\epsilon) > 0$ such that:*

$$\mathbb{P}\left(\sup_{x \in \mathbb{R}^n} \left\{ \left| \hat{b}_1^N(x) - \hat{b}_2^N(x) \right| \right\} > \epsilon \right)$$

$$\leq \Big[C_3(\epsilon)(N+1)n + C_4(\epsilon)\Big] \cdot e^{-N \cdot C_5(\epsilon)}.$$

For sufficiently large N, the factor $N+1$ can be replaced by the constant 2.

Proof. The proof is contained in Appendix B. □

Remark 3.4.1. *If H_0 does not hold, i.e., $\delta := \sup_{x\in\mathbb{R}^n} |b_1(x) - b_2(x)| > 0$, then the same logical arguments used to establish Theorem 3.4.1 can be employed to show that (for the $n = 1$ case), for each $\epsilon > 0$:*

$$\mathbb{P}\left(\sup_{x\in\mathbb{R}^n}\left\{\left|\hat{b}_1^N(x) - \hat{b}_2^N(x)\right|\right\} \in (\delta-\epsilon, \delta+\epsilon)\right)$$
$$\leq C_1(\epsilon)\cdot e^{-N\cdot C_2(\epsilon)},$$

where $C_1(\epsilon)$, $C_2(\epsilon) > 0$ are the same continuous, positive functions given above. That is, as $N\rangle\infty$, the gap between $b_1^N(x)$ and $b_2^N(x)$ approaches δ exponentially. The $n > 1$ case follows analogously from the multivariate version of the Dvoretsky-Kiefer-Wolfowitz (DKW) inequality [144].

3.5 Results

Here, we illustrate the numerical performance of our proposed method on simulated and real-world traffic data, and its efficacy over baseline aggregation methods of concatenating data points along a single, fixed time t. We note that in the experiments on the real-world dataset collected from the Caltrans PeMS system, the data collected is time-ordered. Code containing the datasets and experiments is available at the following link:

`https://github.com/kkulk/causality`.

Simulated Data

In our first set of experiments, we construct synthetic data for single- and multi-link traffic networks. For the single-link network, we use the following dynamics. For each $t \in [T_h]$:

$$x[t+1] = (1-\mu(A[t]))\cdot x(t) + \mu(A[t])\cdot u[t] + w[t],$$
$$A[t+1] \sim \mathcal{P}(x(t))$$

where $x(t) \in \mathbb{R}$ denotes the traffic flow at time t, $A[t] \in \{0,1\}$ is the Boolean random variable that indicates whether or not an accident has occurred at time t, $\mu(A[t]) > 0$ describes the fraction of traffic flow departing the link, $u[t] \in \mathbb{R}$ denotes the total input traffic flow, $w[t] \in \mathbb{R}$ is a zero-mean noise term, and T_h is the finite time horizon. Here, we set $T_h = 500$, $\mu(0) = 0.3$, $\mu(1) = 0.2$, $u(t) = 100$ for each $t \in [T_h]$, and draw $w(t)$ i.i.d. from the continuous uniform distribution on $(-10, 10)$. We create datasets corresponding to the null and alternative hypotheses. For the null hypothesis, we fix the distribution of $x(t)$ to be Bernoulli(0.01), regardless of the value of $x(t)$. This simulates a scenario where the likelihood of an accident occurring has no dependence on traffic flow. For the alternative hypothesis, we set the distribution of $x(t)$ to be Bernoulli(0.01) when $x(t) < 109$ and Bernoulli(0.10) when $x(t) \geq 109$. This represents a scenario where higher traffic loads increase the likelihood that an accident occurs.

To contrast the performance of our algorithm with the baseline, we compute the following quantities from datasets of independent trajectories corresponding to H_0 and H_1, in accordance with Proposition 3.3.2 and Theorem 3.4.1:

- **For our method**—We compute the empirical estimates $\hat{b}_1^N(x)$ and $\hat{b}_2^N(x)$ of the functions $b_1(x)$ and $b_2(x)$ as functions of x (Figure 3.1), and the maximum CDF gap $\sup_{x \in \mathbb{R}^n} |\hat{b}_1^N(x) - \hat{b}_2^N(x)|$ as functions of N (Figure 3.2).

- **For the baseline method**—We compute the empirical estimates of the CDFs of $X_{t-1}|T = t$ and X_{t-1}, with t fixed at 1, as functions of x (Figure 3.1), and the corresponding maximum CDF gap as functions of N (Figure 3.2). Note that for $N < 500$, it is difficult to obtain the CDF of $X_{t-1}|T = t$, due to rarity of the event at any given time.

Figures 3.1 and 3.2 show that, compared to the baseline, our approach distinguishes between the null and alternative hypotheses from a far smaller dataset. This illustrates that our method, compared to the baseline, distinguishes the dependence between the occurrence of a rare event and the state values immediately preceding the event more efficiently.

Appendix B contains further empirical results on synthetic datasets for multi-link networks.

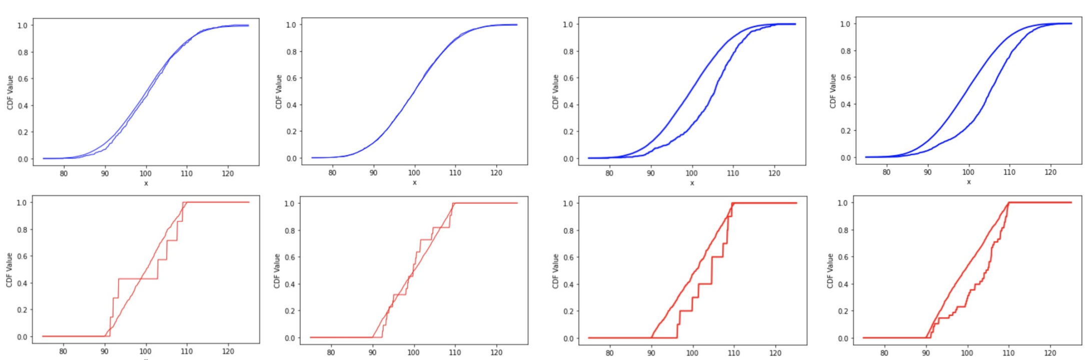


Figure 3.1: (Top) From left to right, $b_1(x)$ and $b_2(x)$ vs. x plots for $(H_0, N = 500)$, $(H_0, N = 2000)$, $(H_1, N = 500)$, and $(H_1, N = 2000)$. (Bottom) From left to right, empirical CDFs for $X_{t-1}|T = t$ and X_{t-1} with $t = 1$, in the same order of hypothesis and N values.

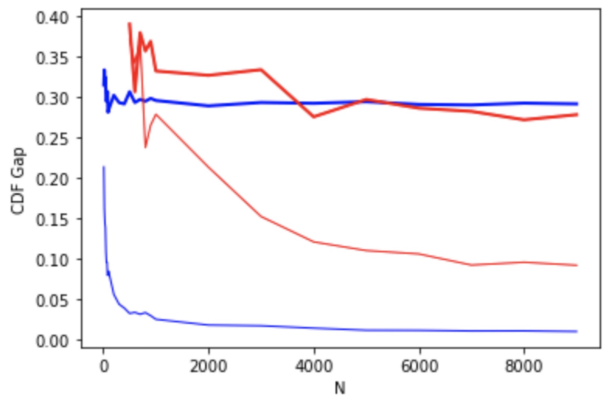


Figure 3.2: CDF gap between vs. N. Red and blue correspond to the baseline and our method, respectively, while thick and thin lines correspond to the null and alternative hypotheses, respectively. Our approach (thin blue curve) correctly identifies the null hypothesis dataset with a relatively small number of samples, while the baseline aggregation method fails to do so (thin red curve).

Caltrans PeMS Dataset

We also demonstrate the efficacy of Algorithm 1 on real traffic flow and incident data collected from the publicly available Caltrans Performance Measurement System (PeMS) dataset [185]. PeMS uses loop detectors placed on freeways to collect flow, speed, and other traffic condition information, and overlays this with incident reports. We consider daily traffic flow data (# vehicles / time) collected from January to August 2022 from 6 A.M. to 2 P.M., at 5-minute intervals, on various bridges in the San Francisco Bay Area: San Mateo-Hayward, San Francisco - Oakland, and Richmond - San Rafael. That is, we consider single link networks connecting a source and destination with the continuous variables $X_t \in \mathbb{R}_+$ corresponding to average flows on the link. Correspondingly, we use incident data collected on these bridges by PeMS in the same time interval from the California Highway Patrol (CHP).

Data Collection We treat each day as an independent trajectory of the traffic flows on every bridge. The PeMS dataset contains flows collected from dual loop detectors placed along the bridges. For each time between 6 A.M. and 2 P.M., we average the flow data recorded by loop detectors on each bridge to obtain the state variable X_t for time t. Mathematically, we define $X_t := \frac{1}{|I|} \sum_{i=1}^{|I|} X_t^i$, where I denotes the set of loop detectors on a single link, and X_t^i denotes the flow measured by detector $i \in I$ at time t. We exclude from our analysis any trajectory on which there was no incident for the entire day, since such trajectories do not contain data relevant to our problem of interest.

Results In Table 1, we enumerate the sample size N and test statistic $\sup_{x\in\mathbb{R}} |\hat{b}_1^N(x) - \hat{b}_2^N(x)|$ for the six traffic links (three bridges, each with two directions of traffic flow). Note the substantial difference in the CDF gap (of nearly 0.178) for the Richmond-San Rafael Bridge, East, compared to all other links, indicating that the flows on this link are particularly causally linked to the first time of incident formation. Further, the San Francisco-Oakland Bay Bridge, East, also has a higher CDF gap (0.081) relative to the West direction, and relative to the other bridges. These gaps are visible in the CDF plots in Figures 3(e) and 3(c), respectively.

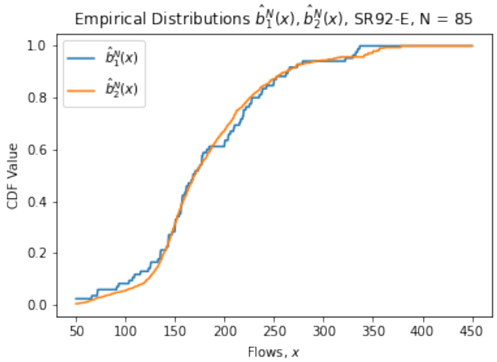
Link	N	$\sup_{x\in\mathbb{R}} \|\hat{b}_1^N(x) - \hat{b}_2^N(x)\|$
San Mateo-Hayward Bridge, East (SR92-E)	85	0.053
San Mateo-Hayward Bridge, West (SR92-W)	116	0.039
San Francisco - Oakland Bay Bridge, East (I80-E)	116	0.081
San Francisco -Oakland Bay Bridge, West (I80-W)	112	0.042
Richmond - San Rafael Bridge, East (I580-E)	45	0.178
Richmond - San Rafael Bridge, West (I580-W)	94	0.048

Table 3.1: CDF gap, $\sup_{x\in\mathbb{R}^n} |\hat{b}_1^N(x) - \hat{b}_2^N(x)|$ for the six links in the San Francisco Bay Area.

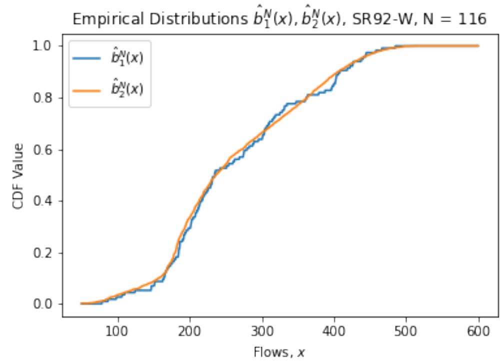
3.6 Discussion and Future Work

We present a novel method for identifying causal links between the state evolution of a dynamical system and the onset of an associated rare event. Crucially, we leverage the time-invariance to reorganize data in a manner that better represents occurrences of the rare event. We then formulate a non-parametric statistical independence test to infer causal dependencies between the dynamical states and the rare event. Empirical results on simulated and real-world time-series data indicate that our method outperforms a baseline approach that conducts independence tests only on a single time slice of the original rare events dataset.

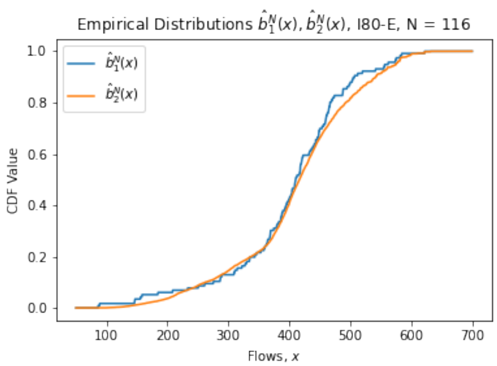
As future work, the causal discovery algorithm presented here may be used to more effectively control the evolution of a dynamical system associated with a rare but consequential event. By establishing causal links between the dynamical state and the rare event, control strategies can be redesigned to maneuver the state away from regions of the state



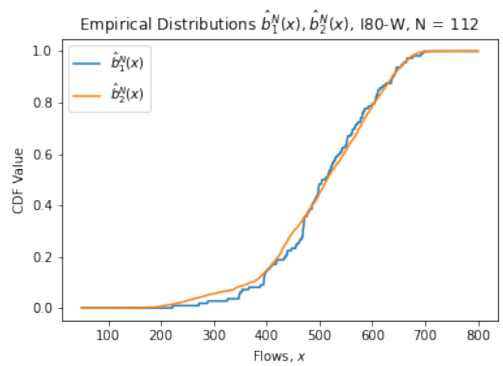
(a) San Mateo-Hayward Bridge, East (SR92-E)



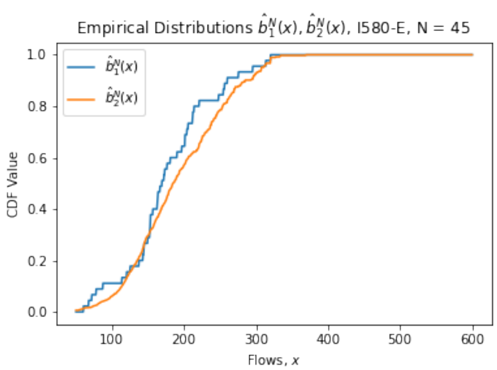
(b) San Mateo-Hayward Bridge, West (SR92-W)



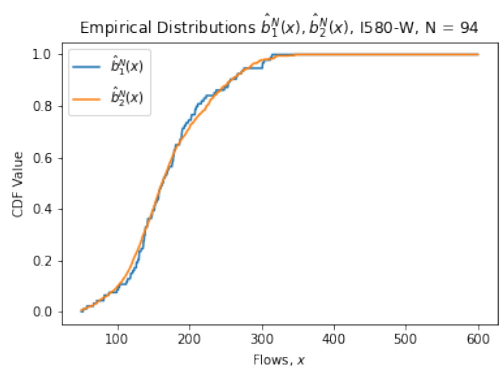
(c) San Francisco - Oakland Bay Bridge, East (I80-E)



(d) San Francisco - Oakland Bay Bridge, West (I80-W)



(e) Richmond-San Rafael Bridge, East (I580-E)



(f) Richmond-San Rafael Bridge, West (I580-W)

Figure 3.3: Empirical CDFs $\hat{b}_1^N(x)$ and $\hat{b}_2^N(x)$ for six bridges in the San Francisco Bay Area.

space where the event occurs more frequently. Important engineering applications include incentive design and flow control methods in the network traffic systems literature, such as dynamic tolling and rerouting. Finally, future work can present more extensive empirical analysis of both the baseline method and our method across different applications.

Chapter 4

Understanding Coalitions Between EV Charging Stations

Apes together strong!

Rise of the Planet of the Apes

This chapter can be found in [106].

4.1 Introduction

The proliferation of electric vehicles (EVs) has brought major changes to road transportation. EVs are playing a significant role in the transition to a sustainable energy-based future and are projected to surpass traditional internal combustion engine-based vehicles in the coming decades [27]. The growth in EVs has led to the genesis of a new industry around building faster and more accessible charging infrastructure [36], comprising several electric vehicles charging companies (EVCCs) such as Tesla, EVgo, and Chargepoint. One major challenge associated with the design of charging infrastructure is that charging stations, especially fast-charging stations, draw a considerable amount of electricity when in operation [183, 6, 63] and may adversely impact the grid infrastructure [5, 117, 62]. The ideal solution to this challenge is to coordinate the demands from all charging stations to distribute the load on the grid [5, 151]. However, this is hard to implement in practice due to the high communication and computational costs of such centralized approaches, and concerns about privacy of the information shared by EV charging stations [36]. Since coordination between all EV charging stations is impractical in most cases, we consider the scenario where only *some* of these stations coordinate and form a *coalition* (refer to Figure 4.1 for a schematic). While it might intuitively seem that some coordination is always better than no coordination, we give a counter-example to show this is not necessarily true. We study the following main question in this chapter:

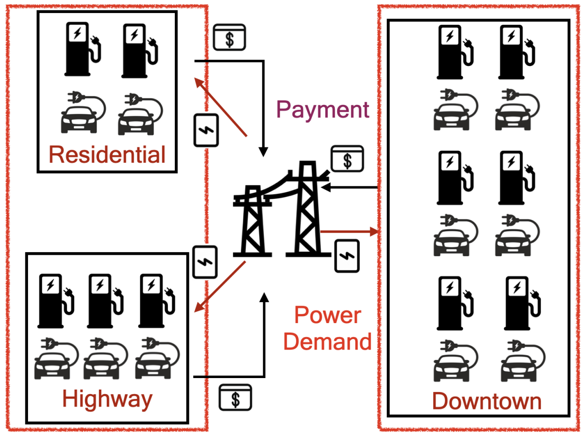


Figure 4.1: A schematic depiction of charging stations located in different areas (e.g. residential, downtown, highway) and owned by different EVCCs. Charging stations that are contained in the same red box are owned by the same company (i.e. form a coalition). Each charging station demands charge from the grid, which determines the prices.

> **Main question (Q):** Could coordination between a few potentially heterogeneous charging stations lead to undesirable consequences?

To answer this question, we model the interaction between charging stations as a non-cooperative aggregative game. Charging stations are strategic agents that draw power from the grid over a finite time window, and have different location-specific charging demands and different sensitivities to deviations from a desired operating demand profile. Each charging station aims to minimize its cost comprising of (i) a payment for power demanded from the grid, and (ii) the deviations from their desired operating charging profile. In our model, we consider that the price per unit of power depends on the total power demanded by all charging stations (resulting in a game theoretic interaction between charging stations). Therefore, in our model, charging stations act as "price-makers", rather than "price-takers" [128, 76, 153, 51]. Finally, some charging stations enter into a coalition (e.g. those owned by a single EVCC) to coordinate their total demand in order to minimize their total cost.

We compare the outcome of the game-theoretic interaction in two scenarios: *(a)* when some charging stations form a coalition $\mathcal{C}$, and *(b)* when each charging station operates independently without any coordination. We differentiate the equilibrium in these two scenarios as $\mathcal{C}-$Nash and Nash equilibrium respectively, and characterize them analytically in Theorem 4.4.1 and Corollary 4.4.3 respectively. In general, we observe that the equilibrium decomposes into two components: a charging profile uniformly distributed across time and a correction term to account for the coalition of charging stations (Theorem 4.4.1). We assess the scenarios *(a)*-*(b)* in terms of three metrics depending on the overall cost experienced by: *(i)* all charging stations (the societal cost); *(ii)* all charging stations within a coalition, and *(iii)* all charging stations outside of a coalition. We identify sufficient conditions on the game parameters under which the formation of a coalition will be worse than the independent operation of charging stations in terms of all three metrics *(i)* − *(iii)*, as presented in Theorems 4.4.4 and 4.4.6. Furthermore, we present numerical examples that satisfy these conditions, demonstrating that coalitions can lead to worse outcomes for all the charging stations, the coalition, and non-coalition charging stations alike. Notably, we find that these outcomes persist even when we relax the assumptions in our theoretical results, highlighting scenarios where conventional intuition about coordinating charging stations may not hold.

4.2 Related Works

Aggregative Games and EV Charging

Several works have proposed studying EV charging as a non-cooperative game [128, 76, 153, 51]. These works mainly study existence, uniqueness, and computation of Nash equilibria of the EV charging game, and analyze these equilibria via measures such as the price of anarchy (PoA) [153]. Our key differentiation from this line of literature is to understand the impact of the formation of a coalition of a subset of EV stations.

Coalitions and Equilibria

Strong Nash equilibrium is an outcome wherein no coalition of agents can collectively deviate from their strategy to improve their utilities, and was first introduced as a concept in [14]. Since this seminal paper, numerous studies have delved into necessary and sufficient conditions for its existence [148], computational properties [77], and its performance across various game classes, often measured by the *k-strong price of anarchy* [68, 61, 7, 66, 15, 39]. This metric quantifies the worst-case welfare loss at strong equilibrium compared to optimal welfare. Unlike strong Nash equilibrium, our notion of $\mathcal{C}-$Nash equilibrium (Definition 4.3.2) is a relaxation and only requires stability with respect to a specific coalition $\mathcal{C}$ and not all possible coalitions. Strong Nash equilibrium may not always exist in our setting but we provide an explicit characterization of $\mathcal{C}$-Nash equilibrium (Theorem 4.4.1).

Users as *Price-Takers*

Prior works have delved into the potential effects of uncoordinated EV charging on the power grid [166, 86]. In response, efforts such as those outlined in [5, 151, 73] have tackled the challenge of devising incentive mechanisms to coordinate small users, often categorized as "price-takers," to shift their charging windows and mitigate grid impacts. However, in contrast to these approaches, our work conceptualizes charging stations, which serve many EVs and thus can aggregate demand, as "price-makers" within electricity markets.

4.3 Model

Notations For any positive integer m, we define $[m] := \{1, 2, ..., m\}$. Consider two matrices $A = (a_{ij})_{i\in[m],j\in[n]} \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times q}$. We define $A \otimes B \in \mathbb{R}^{pm\times qn}$ to be the Kronecker product of matrices A and B. For any finite set X, we define $\mathbf{1}_X \in \mathbb{R}^{|X|}$ to be a vector with all entries to be 1. For any vector $x \in \mathbb{R}^m$ and $p \subseteq [m]$, we use the notation $x_p \in R^{|p|}$ to denote the components of vector x corresponding to p. Additionally, we denote $x_{-p} \in R^{m-|p|}$ to denote the components of vector x corresponding to all entries that are not in p.

Charging stations as strategic entities. Consider a game comprised of $\mathcal{I}$ charging stations, each operating as a strategic entity making decisions on their charging levels throughout a period spanning T units of time. Each station $i \in [N]$ is characterized by a *nominal charging profile* $(\bar{x}_i^t)_{t\in[T]}$, where the total charge demanded is d_i. That is, $\sum_{t\in[T]} \bar{x}_i^t = d_i$. The variation of charge demanded between stations reflects differences in the number of vehicles typically utilizing each charging facility.

The actual charging profile of any station may differ from the nominal charging profile due to the externality imposed by other stations through electricity prices, which we define shortly. We denote x_i^t to be the charge demanded by station i during hour t. Define $X_i = \{x_i \in \mathbb{R}^T : \sum_{t\in[T]} x_i^t = d_i\}$ to be the set of feasible charging profiles for station i [1]. With a slight abuse of notation, we define $x^t \in \mathbb{R}^N$ as the vector of charging profile of all stations at time step $t \in [T]$, and $x = (x_i^t)_{t\in T, i\in\mathcal{I}}$ to be the joint charging profile of all stations.

In our model, we consider reactive prices where the electricity price depends on the total charge demanded. More formally, the cost incurred by any station i under a joint strategy x is given by:

$$c_i(x) = \sum_{t\in[T]} p^t(x)x_i^t + \frac{\mu_i}{2}\|x_i - \bar{x}_i\|^2, \tag{4.1}$$

where (a) $\mu_i > 0$ is the sensitivity parameter of station i; (b) for every $t \in [T]$, $p^t(x)$ denotes the price per unit of electricity when the joint charging profile of all stations is x. It is through this price signal that the charging profiles of other charging stations impose an externality on any station. In what follows we assume that the price function is linear, i.e. $p^t(x) = a^t + b^t\mathbf{1}_N^\top x^t$ for some $a^t, b^t > 0$ ([153, 51]). The parameters a^t represents the price fluctuations due to non-EV demand on the grid. We assume that for all time $t \in [T]$, $b^t = b$ for some positive scalar $b \in \mathbb{R}$.

[1] We do not impose non-negativity constraints on the charging profile. This modeling decision is justified, as charging stations have the capability not only to draw power from the grid but also to inject power into the grid when necessary [104]. Correspondingly, we also do not impose upper bounds on the charging profiles; we assume that there is always enough charge available to meet the charging demand for a given problem instance. It is an interesting direction of future research to impose additional constraints on the set X which align more closely with real-world conditions.

Remark 4.3.1. *The heterogeneity in sensitivity parameters μ_i between different stations can be ascribed to their geographic location (compare Figure 4.1). For instance, a station positioned in a major city's downtown can exhibit heightened sensitivity in meeting its demand requirements compared to one situated within a residential neighborhood.*

Nash equilibrium, defined below, is widely used to characterize the outcome of the interaction in charging games [128, 76, 153, 51].

Definition 4.3.1. *A joint charge profile x^* is a* Nash equilibrium *if $c_i(x_i^*, x_{-i}^*) \le c_i(x_i, x_{-i}^*)$ for every $i \in [N], x_i \in X_i$.*

Coalitions between charging stations. Coalitions between charging stations can be easily facilitated by electric vehicle charging companies who operate multiple charging stations. Every coalition of stations jointly decides their charging profile. For ease of exposition, we consider only single coalition[2]. The goal of the coalition is to choose $x_{\mathcal{C}} \in \prod_{i \in \mathcal{C}} X_i$ that minimizes the coalition's cumulative cost function $c_{\mathcal{C}}(x) = \sum_{i \in \mathcal{C}} c_i(x)$. Next, we introduce a natural extension of Definition 4.3.1 to examine the outcome of interactions in the presence of a coalition.

Definition 4.3.2. *A joint charge profile $x^\dagger \in X$ is a $\mathcal{C}$-*Nash equilibrium *if (i) $\sum_{i \in \mathcal{C}} c_i(x^\dagger) \le \sum_{i \in \mathcal{C}} c_i(x'_{\mathcal{C}}, x^\dagger_{-\mathcal{C}})$ for every $x'_{\mathcal{C}} \in \prod_{j \in \mathcal{C}} X_j$, and (ii) $c_i(x^\dagger) \le c_i(x'_i, x^\dagger_{-i})$ for every $x'_i \in X_i, i \notin \mathcal{C}$.*

When $|\mathcal{C}| = 1$, Definition 4.3.1 and 4.3.2 are equivalent. For the rest of this chapter, we denote $\mathcal{C}-$Nash equilibrium by $x^\dagger$ and Nash equilibrium when charging stations act autonomously by x^*.

4.4 Results

In Section 4.4, we analytically characterize the $\mathcal{C}-$Nash equilibrium in terms of game parameters. Next, we derive sufficient conditions when the Nash equilibrium is preferred over $\mathcal{C}-$Nash equilibrium, in terms of the overall cost experienced by all charging stations, by charging stations within the coalition, and by charging stations outside the coalition. Finally, we provide numerical instances where the coalition performs worse than the Nash equilibrium.

Analytical Characterization of $\mathcal{C}-$Nash equilibrium

Theorem 4.4.1. *For any arbitrary coalition $\mathcal{C} \subset [N]$, when $b > 0$ and $\mu_i > 0\ \forall i \in \mathcal{I}$, the $\mathcal{C}-$Nash equilibrium exists, is unique, and takes the following form*

$$x^{\dagger t} = \frac{d}{T} + \sum_{t' \in [T]} \frac{T\delta^{tt'} - 1}{T} \left(b(\mathbf{1}_{\mathcal{I}} \mathbf{1}_{\mathcal{I}}^\top + \boldsymbol{C}) + \mu \right)^{-1} \cdot$$

[2]The results in this chapter can be extended to encompass scenarios involving multiple coalitions.

$$\cdot \left(\mu\bar{x}^{t'} - a^{t'}\mathbf{1}_{\mathcal{I}}\right), \tag{4.2}$$

where $\delta^{tt'}$ is the Kronecker delta function ($\delta^{tt'} = 1$ when $t = t'$ and is 0 otherwise), $\mu = \textit{diag}([\mu_1, \cdots, \mu_{\mathcal{I}}])$ and $\boldsymbol{C} = \begin{pmatrix} \mathbf{1}_{\mathcal{C}}\mathbf{1}_{\mathcal{C}}^\top & 0 \\ 0 & I_{\mathcal{I}\setminus\mathcal{C}} \end{pmatrix}$.

Proof. Before presenting the proof, we define some useful notation. Define $A = [a^1, a^2, ..., a^T]^\top$ and the concatenated vector of charging profiles $x^\dagger := [x_1^{1^\dagger}, \cdots, x_{\mathcal{I}}^{1^\dagger}, \cdots\cdots, x_1^{t^\dagger}, \cdots, x_{\mathcal{I}}^{t^\dagger}]^\top$.

First, we show that for any coalition $\mathcal{C} \subset [N]$, the resulting game is a strongly monotone game. To ensure this it is sufficient to verify that (i) the strategy set is convex, and (ii) the game Jacobian is positive definite [64]. The convexity of strategy sets holds because the strategy sets X_i are simplices. Next, to compute the game Jacobian, we define $\mathcal{G}_i(x) = \frac{\partial c_{\mathcal{C}}(x)}{\partial x_i}$ if $i \in \mathcal{C}$, and $\mathcal{G}_i(x) = \frac{\partial c_i(x)}{\partial x_i}$ if $i \notin \mathcal{C}$.

It can be verified that the game Jacobian, $J(x)$, is $J(x) = \nabla\mathcal{G}(x) = \Theta$, where $\Theta := I_T \otimes \left(b(\mathbf{1}_{\mathcal{I}}\mathbf{1}_{\mathcal{I}}^\top + \mathbf{C}) + \mu\right)$, which is guaranteed to be a positive definite matrix by Lemma 4.4.2. Thus, for any $\mathcal{C}$, the game is strongly monotone, and the equilibrium is unique. We now introduce two optimization problems:

$$\begin{aligned} P_{\mathcal{C}}(x_{-\mathcal{C}}^\dagger) : \min_{x_{\mathcal{C}} \in X_{\mathcal{C}}} \ & \sum_{j\in\mathcal{C}} c_j(x_{\mathcal{C}}, x_{-\mathcal{C}}^\dagger) \\ \text{s.t.} \ & \sum_{t'\in[T]} x_j^{t'} = d_j \quad \forall j \in \mathcal{C}, \\ \forall\, i \in [N]\setminus\mathcal{C},\ P_i(x_{-i}^\dagger) : \min_{x_i\in X_i} \ & c_i(x_i, x_{-i}^\dagger) \\ \text{s.t.} \ & \sum_{t'\in[T]} x_i^{t'} = d_i. \end{aligned}$$

By definition of $\mathcal{C}-$Nash equilibrium, charging stations within $\mathcal{C}$ jointly solve the optimization problem $P_{\mathcal{C}}(x_{-\mathcal{C}}^\dagger)$ when the charging profiles of other stations $x_{-\mathcal{C}}^\dagger$ is known. Similarly, each player $i \in [N]\setminus\mathcal{C}$ solves the optimization problem $P_i(x_{-i}^\dagger)$ when the charging profile of other players $x_{-i}^\dagger$ is known. Note that each of the optimization problems has linear constraints, and hence the Karush–Kuhn–Tucker (KKT) conditions are necessary for optimality. We solve the KKT conditions of all these problems simultaneously to get a unique solution. Since we established that the equilibrium is unique, this unique solution to the KKT conditions is the $\mathcal{C}-$Nash equilibrium.

The KKT conditions can be written in a compact form, where $\lambda = (\lambda_i)_{i\in[N]}$ are the Lagrange multipliers associated with the linear constraint of every charging station:

$$\Theta x^\dagger = (I_T \otimes \mu)\bar{x} - (I_T \otimes \mathbf{1}_{\mathcal{I}})A - (\mathbf{1}_T \otimes I_{\mathcal{I}})\lambda \tag{4.3a}$$

$$(\mathbf{1}_T^\top \otimes I_{\mathcal{I}})x^\dagger = d, \tag{4.3b}$$

where $\Gamma := (\mathbf{1}_T^\top \otimes I_\mathcal{I})\Theta^{-1}(\mathbf{1}_T \otimes I_\mathcal{I})$. We prove in Lemma 4.4.2 that Γ and Θ are invertible. Using this fact and solving (4.3) by eliminating λ, we obtain a closed-form expression for the $\mathcal{C}-$Nash equilibrium in (4.4):

$$x^\dagger = \Psi_1(\Gamma, \Theta)\left((I_T \otimes \mu)\bar{x} - (I_T \otimes \mathbf{1}_\mathcal{I})A\right) + \Psi_2(\Gamma, \Theta)d, \tag{4.4}$$

with $\Psi_1(\Gamma, \Theta) = \left(I_{NT} - \Theta^{-1}(\mathbf{1}_T \otimes I_\mathcal{I})\Gamma^{-1}(\mathbf{1}_T^\top \otimes I_\mathcal{I})\right)\Theta^{-1}$ and $\Psi_2(\Gamma, \Theta) = \Theta^{-1}(\mathbf{1}_T \otimes I_\mathcal{I})\Gamma^{-1}$. Expanding the terms using the time index, it can be checked that equation (4.4) is equivalent to (4.2), completing the proof. □

Next, we present a technical result, used in the proof of Theorem 4.4.1, the proof of which is deferred to Appendix C.

Lemma 4.4.2. *Θ and Γ are positive definite and hence invertible matrices.*

We now make some remarks about Theorem 4.4.1.

Remark 4.4.1. *The closed-form equilibrium solution in (4.2) is comprised of a fixed term and a correction term. The fixed term represents a charging profile that is uniform across time. The second term is the correction to account for the aggregate effects of the coalition, demand and charging preferences of the charging stations, and exogenous non-EV demand. As expected, when the time-dependent factors $a^{t'}$ and $\bar{x}^{t'}$ are constant across time, the correction term is zero and charging stations charge uniformly across time.*

Remark 4.4.2. *The result in Theorem 4.4.1 extends to the setting of multiple (non-overlapping) coalitions $\mathcal{C}_1, \mathcal{C}_2, \ldots \mathcal{C}_K$ by setting*

$$\boldsymbol{C} = \begin{pmatrix} \mathbf{1}_{\mathcal{C}_1}\mathbf{1}_{\mathcal{C}_1}^\top & 0 & \cdots & 0 \\ 0 & \mathbf{1}_{\mathcal{C}_2}\mathbf{1}_{\mathcal{C}_2}^\top & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \mathbf{1}_{\mathcal{C}_K}\mathbf{1}_{\mathcal{C}_K}^\top & \vdots \\ 0 & 0 & 0 & I_{\mathcal{I}\setminus\cup_{i=1}^K \mathcal{C}_i} \end{pmatrix}.$$

In fact, all theoretical results in this article can be directly extended to the multiple coalition case by using the above $\boldsymbol{C}$ matrix. For the sake of clear presentation, we shall always work with $K = 1$ in the following.

We conclude this section by stating the following specialization of Theorem 4.4.1, which characterizes Nash equilibrium.

Corollary 4.4.3. *The Nash equilibrium x^* takes the following form:*

$$x^{*t} = \frac{d}{T} + \sum_{t' \in [T]} \frac{T\delta^{tt'} - 1}{T} \left(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + I_N) + \mu\right)^{-1} \cdot \left(\mu\bar{x}^{t'} - a^{t'}\mathbf{1}_\mathcal{I}\right),$$

where $I_N \in \mathbb{R}^{N \times N}$ is the identity matrix.

The proof of Corollary 4.4.3 follows by setting $\mathcal{C} = \{1\}$ in Theorem 4.4.1.

When is $\mathcal{C}-$Nash equilibrium beneficial?

In this subsection, we study conditions under which the $\mathcal{C}-$Nash equilibrium is preferred to Nash equilibrium and vice-versa. For a charging profile x, define the total cost incurred by a group of charging stations in $\mathcal{S} \subseteq [\mathcal{I}]$ as

$$c_{\mathcal{S}}(x) := \sum_{i \in \mathcal{S}} c_i(x), \quad \forall\ x \in X. \tag{4.5}$$

In order to compare the outcome under Nash equilibrium and $\mathcal{C}-$Nash equilibrium, we use the following metrics:

$$\underbrace{\frac{c_{[\mathcal{I}]}(x^*)}{c_{[\mathcal{I}]}(x^\dagger)}}_{:=\mathsf{M}_{[\mathcal{I}]}}, \quad \underbrace{\frac{c_{\mathcal{C}}(x^*)}{c_{\mathcal{C}}(x^\dagger)}}_{:=\mathsf{M}_{\mathcal{C}}}, \quad \underbrace{\frac{c_{[N]\backslash\mathcal{C}}(x^*)}{c_{[N]\backslash\mathcal{C}}(x^\dagger)}}_{:=\mathsf{M}_{[N]\backslash\mathcal{C}}}, \tag{Eval-Metric}$$

where, recall, x^* is the Nash equilibrium when there is no coalition, and $x^\dagger$ is the $\mathcal{C}$-Nash equilibrium. These metrics are the ratio of the total cost experienced by different groups of players at Nash equilibrium and at $\mathcal{C}-$Nash equilibrium. In particular, $\mathsf{M}_{[N]}$ is this ratio for all players, $\mathsf{M}_{\mathcal{C}}$ is this ratio for players within the coalition $\mathcal{C}$, and $\mathsf{M}_{[N]\backslash\mathcal{C}}$ is this ratio of players outside of the coalition. For any $\mathcal{S} \in \{[N], \mathcal{C}, [N]\backslash\mathcal{C}\}$, if $\mathsf{M}_{\mathcal{S}} < 1$, the Nash equilibrium is preferred. Otherwise the $\mathcal{C}-$Nash equilibrium is preferred by the set of players $\mathcal{S}$. Next, we theoretically characterize these metrics in two cases:

Case A: Exogeneous price fluctuations a^t are uniform across time

Here, we specialize to the case when $a^t = a\ \forall t \in [T]$ for some $a \in \mathbb{R}$. This can happen when the price is influenced by a constant non-EV demand throughout the day. Under this setting, the equilibria $x^\dagger$ and x^* are given below.

Proposition 8. *Suppose $a^t = a^{t'}$ for all $t, t' \in [T]$. Then for every $t \in [T]$,*

$$x^{*t} = \frac{d}{T} + \sum_{t' \in [T]} \frac{T\delta^{tt'} - 1}{T} \left(b(\mathbf{1}_{\mathcal{I}}\mathbf{1}_{\mathcal{I}}^\top + I_N) + \mu\right)^{-1} \mu \bar{x}^{t'},$$

$$x^{\dagger t} = \frac{d}{T} + \sum_{t' \in [T]} \frac{T\delta^{tt'} - 1}{T} \left(b(\mathbf{1}_{\mathcal{I}}\mathbf{1}_{\mathcal{I}}^\top + C) + \mu\right)^{-1} \mu \bar{x}^{t'}.$$

Further, consider the scenario when all the charging stations prepare for similar peak and low demand hours, and use charging rate recommendations from EV manufacturers to predict their demand requirements. That is, they desire similar demand profiles (up to a constant factor) across the day. This scenario is captured in Assumption 4.4.1.

Assumption 4.4.1. *The desired demand of each charging station $i \in [N]$ at time step $t \in [T]$ is $\bar{x}_i^t = d_i \alpha^t$, where $\sum_{t \in [T]} \alpha^t = 1$ and $\alpha^t \geq 0$ for all $t \in [T]$.*

Next, we delineate conditions under which the formation of a coalition is worse than the independent operation of charging stations in terms of verifiable conditions on game parameters.

Theorem 4.4.4. *Suppose Assumption 4.4.1 holds and $a^t = a^{t'}$ for all $t, t' \in [T]$. Any group $\mathcal{S} \subseteq [N]$ incurs a lower cost in Nash equilibrium compared to $\mathcal{C}-$Nash equilibrium if and only if*

$$\sum_{i\in\mathcal{S}} f_i^{\dagger}(\mu, b, d) - f_i^{*}(\mu, b, d) \geq 0, \tag{4.6}$$

where for every $i \in [N]$, $f_i^{\dagger}(\mu, b, d) := \Delta^{\dagger}\Delta_i^{\dagger} + \frac{\mu_i(\Delta_i^{\dagger}-d_i)^2}{2b}$, $f_i^{*}(\mu, b, d) := \Delta^{*}\Delta_i^{*} + \frac{\mu_i(\Delta_i^{*}-d_i)^2}{2b}$, $\Delta_i^{\dagger} := \big((b(1_N 1_N^{\top} + \boldsymbol{C}) + \mu)^{-1}\mu d\big)_i$, and $\Delta_i^{*} := \big((b(1_N 1_N^{\top} + I_N) + \mu)^{-1}\mu d\big)_i$. Additionally, $\Delta^{\dagger} := \sum_{i\in[N]} \Delta_i^{\dagger}$ and $\Delta^{*} := \sum_{i\in[N]} \Delta_i^{*}$.

Proof. For a subset $\mathcal{S}$ of charging stations, the total cost under Nash equilibrium is lower than $\mathcal{C}-$Nash equilibrium if and only if

$$\sum_{i\in\mathcal{S}} c_i(x^*) \leq \sum_{i\in\mathcal{S}} c_i(x^{\dagger}). \tag{4.7}$$

In the rest of the proof, we calculate these costs in terms of the game parameters. For every $t \in [T]$, define $F^t := T\alpha^t - \sum_{t'\in[T]} \alpha^{t'}$. Since $\sum_{t\in[T]} \alpha^t = 1$, it holds that $F^t = T\alpha^t - 1$ and $\sum_{t\in[T]} F^t = 0$. Using Assumption 4.4.1 along with Proposition 8, we get the following for $\mathcal{C}-$Nash equilibrium:

$$\begin{aligned} x_i^{\dagger t} = \frac{d_i}{T} + \frac{F^t}{T}\Delta_i^{\dagger}, \qquad \mathbf{1}_N^{\top} x^{\dagger t} &= \frac{D}{T} + \frac{F^t}{T}\Delta^{\dagger} \\ x_i^{\dagger t} - \bar{x}_i^t &= \frac{F^t}{T}(\Delta_i^{\dagger} - d_i), \end{aligned} \tag{4.8}$$

where $D = \sum_{i\in[N]} d_i$. Using (4.1) and (4.8) and $\sum_{t\in[T]} F^t = 0$, the cost of station i at $\mathcal{C}-$Nash equilibrium is:

$$\begin{aligned} c_i(x^{\dagger}) &= a d_i + \left(b \sum_{t\in[T]} (1_N^{\top} x^{t\dagger}) x_i^{t\dagger} \right) + \frac{\mu_i}{2}\|x_i^{\dagger} - \bar{x}_i\|^2 \\ &= a d_i + \frac{b}{T^2}\left(TDd_i + \sum_{t\in[T]} (F^t)^2 \Delta^{\dagger}\Delta_i^{\dagger} \right) \\ &\quad + \frac{\mu_i}{2T^2} \sum_{t\in[T]} (F^t)^2 (\Delta_i^{*} - d_i)^2. \end{aligned}$$

Consequently, the total cost of charging stations in $\mathcal{S}$ is:

$$\sum_{i\in\mathcal{S}} c_i(x^\dagger) = \left(aD_\mathcal{S} + \frac{bDD_\mathcal{S}}{T}\right) + \frac{\mathcal{F}}{T^2}\left(\sum_{i\in\mathcal{S}} b\Delta^\dagger\Delta_i^\dagger + \frac{\mu^i}{2}(\Delta_i^\dagger - d_i)^2\right),$$

where $D_\mathcal{S} := \sum_{i\in\mathcal{S}} d_i$, $\mathcal{F} := \sum_{t\in[T]}(F^t)^2$. Analogously, we can also compute the total cost at Nash equilibrium. Using these results, we conclude (4.7) is equivalent to

$$\begin{aligned} &aD_\mathcal{S} + \frac{bDD_\mathcal{S}}{T} + \frac{\mathcal{F}}{T^2}\left(\sum_{i\in\mathcal{S}} b\Delta^\dagger\Delta_i^\dagger + \frac{\mu^i}{2}(\Delta_i^\dagger - d_i)^2\right) \\ &\geq aD_\mathcal{S} + \frac{bDD_\mathcal{S}}{T} + \frac{\mathcal{F}}{T^2}\left(\sum_{i\in\mathcal{S}} b\Delta^*\Delta_i^* + \frac{\mu^i}{2}(\Delta_i^* - d_i)^2\right) \\ &\iff \sum_{i\in\mathcal{S}} f_i^\dagger(\mu,b,d) - f_i^*(\mu,b,d) \geq 0. \end{aligned}$$

□

Remark 4.4.3. *Interestingly, the condition* (4.6) *in Theorem 4.4.4 is independent of* T, α^t, a *and depends only on* μ *and* d.

Corollary 4.4.5. *When* $c_i(x^\dagger), c_i(x^*) \geq 0\ \forall i \in [N]$, *Theorem 4.4.4 can be used to analyze (Eval-Metric) as follows. For any* $\mathcal{S} \in \{[N], \mathcal{C}, [\mathcal{N}]\backslash\mathcal{C}\}$, $\mathsf{M}_\mathcal{S} \leq 1$ *if and only if* $\sum_{i\in\mathcal{S}} f_i^\dagger(\mu,b,d) - f_i^*(\mu,b,d) \geq 0$.

Case B: The desired demand $\bar{x}^t$ is uniform across time

In this subsection, we analyze the setting when $\bar{x}^t$ is uniform in t. This denotes a uniform spread of desired demand by the stations. This resembles the nominal charging profile when the EV adoption will reach a critical mass, when each charging station observes a constant flow of EVs. In this setting, when averaged over any time window, the nominal charge is uniformly spread.

Proposition 9. *If* $\bar{x}^t = d/T, \quad \forall\ t \in [T]$ *then*

$$\begin{aligned} x^{*t} &= \frac{d}{T} - \sum_{t'\in[T]} \frac{T\delta^{tt'} - 1}{T}\left(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + I_N) + \mu\right)^{-1}\mathbf{1}_N a^{t'}, \\ x^{\dagger t} &= \frac{d}{T} - \sum_{t'\in[T]} \frac{T\delta^{tt'} - 1}{T}\left(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + C) + \mu\right)^{-1}\mathbf{1}_N a^{t'}. \end{aligned}$$

Theorem 4.4.6. *Suppose $\bar{x}^t = d/T$, for every $t \in [T]$. Any group $\mathcal{S} \subseteq [N]$ incurs a lower cost in Nash equilibrium when compared to $\mathcal{C}-$Nash equilibrium if and only if $g_{\mathcal{S}}^{\dagger}(\mu, b) - g_{\mathcal{S}}^{*}(\mu, b) \geq (\Gamma^{\dagger} - \Gamma^{*})h(A)$, where for every $i \in [N]$,*

$$
\begin{aligned}
g_{\mathcal{S}}^{\dagger}(\mu, b) &:= \frac{b}{T} \sum_{i \in \mathcal{S}} \Gamma^{\dagger} \Gamma_i^{\dagger} + \frac{\mu^i}{2T} (\Gamma_i^{\dagger})^2 \\
g_{\mathcal{S}}^{*}(\mu, b) &:= \frac{b}{T} \sum_{i \in \mathcal{S}} \Gamma^{*} \Gamma_i^{*} + \frac{\mu^i}{2T} (\Gamma_i^{*})^2 \\
h(A) &:= \frac{\sum_{t,t' \in [T]} a^t a^{t'} \left(T\delta^{tt'} - 1\right)}{\sum_{t \in [T]} \left(\sum_{t' \in [T]} a^{t'} \left(T\delta^{tt'} - 1\right)\right)^2}, \\
\Gamma_i^{*} &:= \left((b(1_N 1_N^\top + I_N) + \mu)^{-1} \mathbf{1}_N\right)_i, \\
\Gamma_i^{\dagger} &:= \left(\left(b(\mathbf{1}_{\mathcal{I}} \mathbf{1}_{\mathcal{I}}^\top + C) + \mu\right)^{-1} \mathbf{1}_{\mathcal{I}}\right)_i.
\end{aligned}
$$

Additionally, $\Gamma^{*} := \sum_{i \in [N]} \Gamma_i^{*}$ and $\Gamma^{\dagger} := \sum_{i \in [N]} \Gamma_i^{\dagger}$.

Proof. The proof is analogous to that of Theorem 4.4.4 and is deferred to Appendix C. □

Remark 4.4.4. *The condition in Theorem 4.4.6 does not depend on the heterogeneity of demand of charging stations d.*

Corollary 4.4.7. *When $c_i(x^{\dagger}), c_i(x^{*}) \geq 0 \ \forall i \in [N]$, Theorem 4.4.6 can be used to analyze (Eval-Metric) as follows.*

$$
\begin{aligned}
\mathsf{M}_{\mathcal{S}} \leq 1 \iff g_{\mathcal{S}}^{\dagger}(\mu, b) - g_{\mathcal{S}}^{*}(\mu, b) &\geq (\Gamma^{\dagger} - \Gamma^{*})h(A) \\
&\forall \mathcal{S} \in \{[N], \mathcal{C}, [\mathcal{N}] \backslash \mathcal{C}\}.
\end{aligned}
$$

Coalitions may not be always beneficial.

In this subsection, we construct instances where the formation of a coalition is not beneficial, which satisfy the setup discussed in the previous section. We set $N = 5, T = 10, b = 0.5$ and consider that each station can be of two types: Type H or Type L. For the purpose of this example, we consider that all stations in a coalition are of Type L and all stations outside the coalition are of Type H. A station $i \in [N]$ is said to be Type H if $d_i = 5$ and $\mu_i = 1$, and of Type L if $d_i = 1$ and $\mu_i = 0.1$. Additionally, we set $\alpha^t = \eta_1$ for $t \leq T/2$, and η_2 otherwise. Furthermore, we set $a_t = 0.5 + \delta\nu_t$ where $\nu_t \sim \mathsf{Unif}([0, 1])$. We study the impact of the size of the coalition on various metrics presented in (Eval-Metric). In Figure 4.2, we study the case $\eta_1 = 0.4/T, \eta_2 = 1.6/T$ and $\delta = 0$, which is aligned with the setup considered in Theorem 4.4.4. Meanwhile, in Figure 4.3, we set $\eta_1 = 1/T, \eta_2 = 1/T$ and $\delta = 0.4$, which is aligned with the setup considered in Theorem 4.4.6. From these figures we conclude that

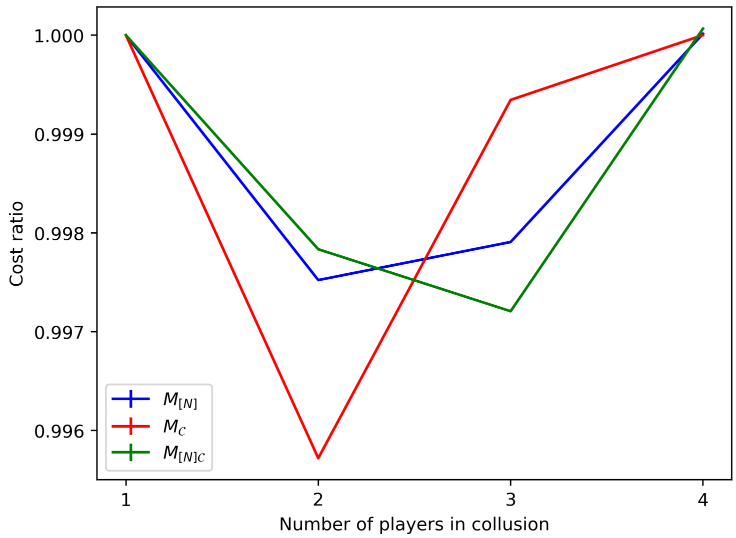


Figure 4.2: Setting $\eta_1 = 0.4/T, \eta_2 = 1.6/T$ and $\delta = 0$

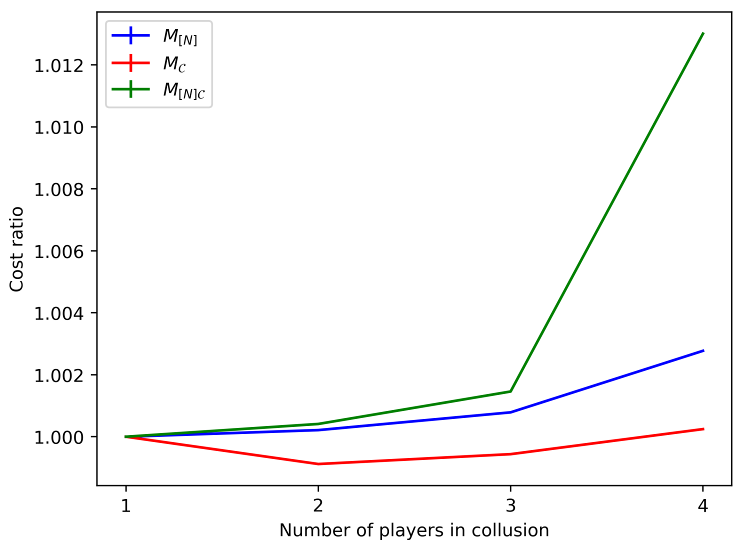


Figure 4.3: Setting $\eta_1 = 1/T, \eta_2 = 1/T$ and $\delta = 0.4$

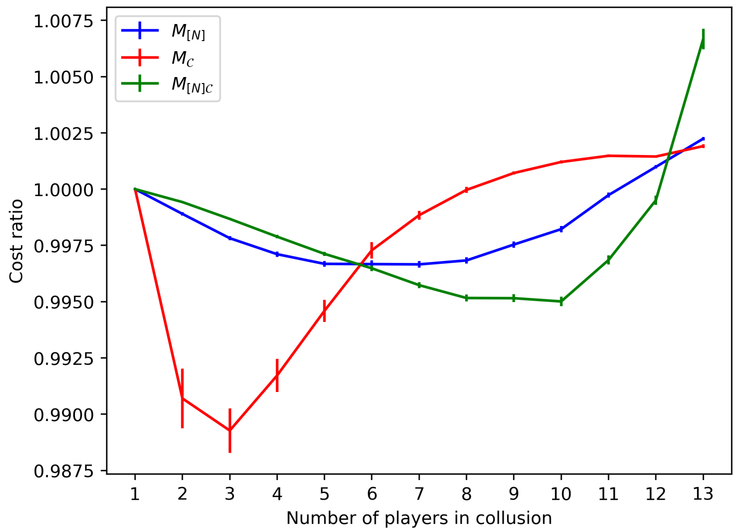


Figure 4.4: Comparison between Nash equilibrium and $\mathcal{C}-$Nash equilibrium with respect to (Eval-Metric) under different sized coalitions.

there exist coalitions, where from the coalition's perspective (i.e. $\mathsf{M}_{\mathcal{C}}$) the outcome under Nash equilibrium is preferred to the outcome under coordination. Furthermore, in Figure 4.2, we find that there exist instances when the $\mathcal{C}-$Nash equilibrium is not preferred under every evaluation metric. Finally, in Figure 4.3, we find that the formation of a coalition not only adversely impacts the coalition but also provides an advantage to stations outside the coalition.

4.5 Numerical Results

In this section we expand on the experimental results from Section 4.4 by relaxing the conditions imposed on game parameters. The code to generate the figures in this section is available at

`https://github.com/kkulk/coalition-ev`.

Impact of simultaneous variations in a^t and α^t. In Figure 4.4, we study the impact of the size of the coalition in terms of (Eval-Metric). In contrast to Section 4.4, we consider both a^t and α^t to be non-uniform and randomly assign stations to be of Type H with probability 0.2 if they are within the coalition. All stations outside of the coalition are assigned to be of Type L. We find that if the size of coalition is small then Nash equilibrium turns out to be favorable along all metrics (Eval-Metric). However, as the size of the coalition increases, it may be beneficial to form a coalition.

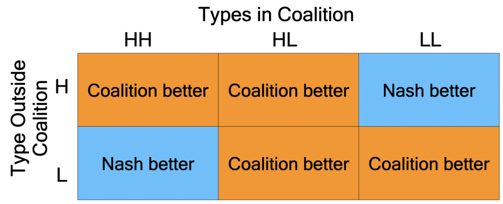


Figure 4.5: The phrase "Coalition better" (resp. "Nash better") implies $\mathsf{M}_{[N]} > 1$ (resp. $\mathsf{M}_{[N]} < 1$).

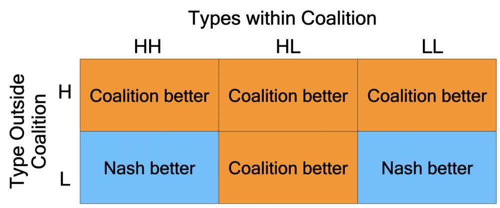


Figure 4.6: The phrase "Coalition better" (resp. "Nash better") implies $\mathsf{M}_{\mathcal{C}} > 1$ (resp. $\mathsf{M}_{\mathcal{C}} < 1$).

Impact of the composition of coalitions. Here we show that not only the size of the coalition, but also the composition of the coalition, plays an important role in deciding whether the coalition is beneficial. To illustrate this point, we examine an example with $N = 3, T = 10, b = 1$ and $a^t = 10$ for all $t \in [T]$. We posit a scenario where the first two stations form a coalition. Each station is one of two types, namely Type H or Type L. We call a station $i \in [N]$ to be of Type H if $d_i = 5$ and $\mu_i = 5$, and of Type L if $d_i = 1$ and $\mu_i = 1$. In Figures 4.5, 4.6, and 4.7, we present a comparative analysis, evaluating how different coalitions perform in terms of metric presented in (Eval-Metric). We find under some circumstances $\mathsf{M}_{\mathcal{S}} > 1$ for all $\mathcal{S} \in \{[N], \mathcal{C}, [N]\backslash[\mathcal{C}]\}$. For instance, this is the case when the coalition is comprised of atleast one Type H station and the station outside the coalition is of Type H. On the contrary, perhaps surprisingly, there are also instances when $\mathsf{M}_{\mathcal{S}} < 1$ for all $\mathcal{S} \in \{[N], \mathcal{C}, [N]\backslash[\mathcal{C}]\}$. For instance, this is the case when the stations within a coalition are of Type H and the station outside the coalition is of Type L.

4.6 Discussion and Future Work

In this work, we initiate a study to understand the impact of coalitions between charging stations as charging infrastructure continues to grow in the coming years. As charging stations draw a substantial amount of electricity from the grid, they will be "price-makers" on the electricity grid. More often than not, multiple charging stations are operated by same electric vehicle charging company (EVCC), which could facilitate coordination between

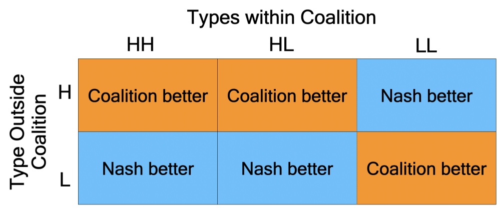


Figure 4.7: The phrase "Coalition better" (resp. "Nash better") implies $\mathsf{M}_{[N]\setminus\mathcal{C}} > 1$ (resp. $\mathsf{M}_{[N]\setminus\mathcal{C}} < 1$).

charging stations. In this work, we analytically characterize the equilibrium outcome in the presence of coalitions. Our analysis hints at potential losses encountered by EVCCs if they coordinate all charging stations owned by them in the presence of heterogeneity in charging demand and user preferences.

There are several interesting questions for future research in understanding the impact of coalitions between charging stations. In the current model we assume the overall demand of a charging station is fixed, but this demand could be affected by electricity prices [117, 62, 193]. We also assume the price functions are linear; future work may extend this to generic nonlinear functions. Furthermore, there are additional operational constraints such as limited capacity of the energy infrastructure and bounds on charging rates of EVs that could be accounted for while computing the equilibrium.

Chapter 5

Redesigning Congestion Pricing

> Defeating traffic is the ultimate boss battle. Even the most powerful humans in the world cannot defeat traffic.
>
> Elon Musk

5.1 Introduction

Congestion pricing is an incentive mechanism for effective utilization of road infrastructure among travelers who selfishly seek to minimize their travel time on the network. Widely adopted in many major cities, both theoretical [191] and empirical [45, 163, 159, 58] studies have shown that congestion pricing can reduce traffic congestion and improve air quality [127, 116, 196, 89]. The revenue generated from congestion pricing is often reinvested to improve the road infrastructure and public transit [82, 176]. Despite these benefits, the implementation of congestion pricing often faces challenges, and one of the primary concerns is its disproportional impact on travelers who have a heterogeneous value-of-time to travel on the network (for example, due to variations in income) [54, 78]. Some travelers often have limited access to alternative transportation options, and congestion fees may add additional financial burden.

In this work we present a principled approach to compute congestion pricing schemes that incorporate both *(i)* the *efficiency* objective of minimizing the total travel time on the network, and *(ii)* the *distribution-welfare* objective, where the distribution is assessed in terms of the maximum disparity in relative change in travel costs experienced by different traveler populations following the implementation of tolls, compared to a scenario with no tolls, and welfare is assessed as the average relative change in travel costs experienced by travelers across all types following the implementation of tolls, compared to a scenario with no tolls. Crucially, we seek to understand the set of efficiency-inducing tolls that also minimize the disparity in costs across different traveler types.

We consider a non-atomic routing game, where travelers make routing decisions based on the travel time of each route plus the monetary cost that includes tolls and gas prices. The monetary cost is adjusted by the travelers' value-of-time— the amount of money a traveler is willing to pay to save a unit of time. Our game has a finite number of traveler populations, each with a heterogeneous value-of-time. Following the result from [88], the equilibrium flow in our game is unique, and can be computed by solving a convex optimization problem. Moreover, the congestion minimizing edge flow vector (i.e. the edge flow vector that minimizes the total travel time) is unique.

We propose four kinds of congestion pricing schemes that differ in terms of whether (a) tolls price different types differently, and (b) whether tolls can be set on all edges or only on a subset of edges. In particular, the four congestion pricing schemes are: (i) *homogeneous pricing scheme with no support constraints*, denoted by hom, where all populations are charged with the same tolls and all edges are allowed to be tolled; (ii) *heterogeneous pricing scheme with no support constraints*, denoted by het, where populations are charged with differentiated toll prices based on their types and all edges can be tolled; (iii) *homogeneous pricing scheme with support constraints*, denoted by hom_sc, where tolls are not differentiated but only a subset of edges can be tolled; (iv) *heterogeneous pricing scheme with support constraints*, denoted by het_sc, where tolls are differentiated and only a subset of edges are tolled.

We compute the tolls in each pricing scheme using a two-step approach. First, we characterize the set of tolls that minimize the total travel time (i.e. efficiency objective). Second, we select a particular toll price in the set of tolls computed in the first step to optimize for an objective that achieves the trade-off between average welfare of all populations and the distribution of costs across different populations. Under the hom and het pricing schemes, the set of tolls that minimize the total travel time (as computed in the first step) can be characterized as the set of solutions of a linear program, and the second step of selecting a particular toll price is also an optimal solution of a linear program (Proposition 5.4.3) and het (Proposition 5.4.4). The two step approach and the linear program formulations build on the study of enforceable equilibrium flows in routing games with heterogeneous populations [70, 103, 191, 91]. On the other hand, under hom_sc and het_sc, direct extensions of the two linear programs to include toll support set constraints are not guaranteed to achieve the efficiency objective. In fact, the problem of designing congestion minimizing pricing schemes with support constraints is known to be NP hard without the consideration of heterogeneous value-of-time [95, 29, 90]. Building on the linear programming-based approaches developed for the pricing schemes without support constraints, we propose a linear programming-based heuristic to compute tolls with support constraints and evaluate their efficiency outcomes in the case study.

We next apply our results to evaluate the performances of the four congestion pricing schemes in the San Francisco Bay Area freeway network. Populations in the San Francisco Bay area exhibit significant income disparities. This is evident from the distribution of

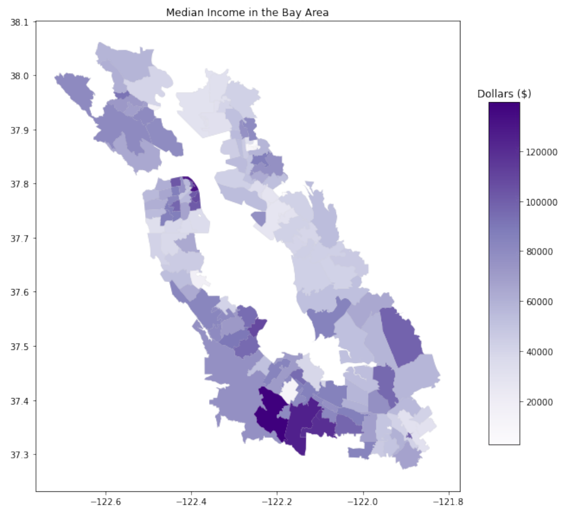


Figure 5.1: Median income.

median annual individual income of each neighborhood as shown in Figure 5.1. Moreover, the area has low public transport coverage and thus, a large fraction of the populations commute via car. We can see in Figure 5.2 that the driving population percentages of most zip codes outside of San Francisco and Oakland are higher than 60%. Moreover, zip codes that are on the east side of the Bay Area have both a high percentage of driving population and a low median individual income. This observation underscores the importance of designing efficient congestion pricing schemes that accounts for the heterogeneity in the willingness-to-pay of different traveler populations.

We model the freeway network in the San Francisco Bay Area as a network with 17 nodes (Figure 5.3). Each node represents a major work or home location for travelers, and the edges represent the primary freeways connecting these locations. Since we differentiate populations based on their value-of-time, which is a latent parameter that cannot be directly estimated from the data, we use the median individual income as a proxy ([13, 87, 188, 184]) to categorize travelers with home at each node into three types of populations with low, middle and high value-of-time, respectively.

Using high-fidelity datasets from Safegraph, the Caltrans Performance Measurement System (PeMS), and the American Community Survey (ACS), we calibrate the latency function of each edge and the demand of each traveler population between each pair of nodes.

The current congestion pricing scheme, denoted by curr , sets a $7 price on each of the bridges in the Bay Area in the east to west direction (Figure 5.3). We compute the four congestion pricing schemes (hom, het, hom_sc, het_sc), and compare the resulting equilibrium routing behavior in comparison to curr and the zero pricing scheme that set no tolls. We summarize our findings below:

(i) Efficiency and Distribution: All four proposed pricing schemes leads to a lower value

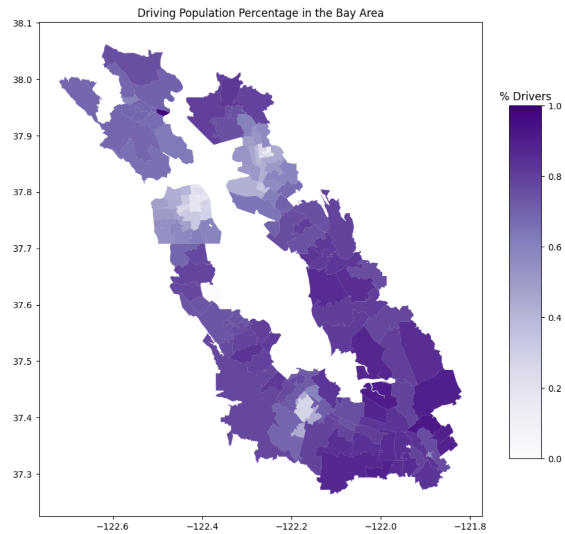


Figure 5.2: Driving population percentage.

of total travel time compared to curr. Surprisingly, curr is also marginally outperformed by zero. This is primarily attributed to the fact that the homogeneous toll price of $7 on all the bridges under curr does not account for the heterogeneous distribution of populations between different home-work locations. We show that hom and het achieve the minimum congestion, as indicated by our theoretical result (Proposition 5.4.2). Additionally, hom_sc and het_sc achieve lower value of total travel time than curr and zero but higher than hom and het. Furthermore, we find that the price of anarchy (POA) – the ratio between the total travel time in equilibrium with no tolls and that of the minimum value total travel time [169] – in our setup is 1.04, which is close to 1. This is likely due to the high total demand of travelers in the Bay area network since the POA always converges to 1 in routing games as the population demand increases [41, 42].

We find that all pricing schemes, except hom, result in lower travel costs for all traveler populations compared to curr. Additionally, our results show that curr is outperformed by all other schemes, except hom, even on the distribution metric.

(ii) Revenue Generation: We observe that the revenue generated by hom is the highest among all schemes, since it charges high tolls to all travelers in order to achieve the minimum congestion. Moreover, the revenues generated by het, hom_sc and het_sc are comparable to curr with het being marginally higher and, hom_sc and het_sc, marginally lower.

The rest of the chapter is organized as follows: Section 5.2 presents an overview of past related works. Section 5.3 presents the model of routing games with heterogeneous populations. Section 5.4 presents computational methods for the four congestion pricing schemes (hom, het, hom_sc, and het_sc). Section 5.5 presents the calibration of the routing game model in the San Francisco Bay Area. Section 5.6 presents the efficiency and distributional evaluations of the proposed pricing schemes and the comparison of their congestion patterns.

5.2 Related Works

The literature on designing congestion pricing schemes can be categorized into two main threads: *first-best* and *second-best*. First-best pricing schemes allow tolls to be placed on every edge of the network. The most popular first-best tolling scheme is marginal cost pricing, which sets the toll price to be the marginal cost created by an additional unit of congestion on each edge [10, 22, 178, 169]. Additionally, an extensive line of research in this thread also focuses on characterizing the set of all congestion-minimizing toll prices (see [191], and references therein). On the other hand, second-best pricing schemes restrict the set of edges that can be tolled. The literature on second-best pricing schemes primarily focuses on formulating the problem as a mathematical program with equilibrium constraints (MPEC) and developing algorithms to approximate the optimal solution (e.g. [192, 35, 69, 108, 110, 123, 156, 186, 112, 56, 102]). The papers [95, 29, 90] studied the problem of characterizing the hardness of the problem of designing second-best tolls. The paper [95] showed that it is NP hard to compute optimal tolls on a subset of edges in general networks and gave a polynomial time algorithm to solve the problem for the parallel link case with affine latency functions. This was extended to allow for non-affine latency functions by [90], and an upper bound on the toll values in [29]. In our setup, hom and het are first-best pricing schemes and hom_sc and het_sc are second-best pricing schemes. We contribute to this line of literature by proposing a multi-step linear programming based approach to compute hom and het that account for the distributional objective and the heterogeneous traveler populations. Our approach also proposes an efficient heuristic to solve hom_sc and het_sc with at most three linear programs instead of iteratively computing the Wardrop equilibrium.

The literature on congestion pricing has mostly focused on homogeneous pricing schemes with a few exceptions. The paper [67] considered tolling schemes that differentiate conventional vehicles from clean energy vehicles. Differentiated tolls are also used in [115, 114, 138] to study mixed autonomy. The paper [37] studied the impact of differentiated tolling in parallel-link networks with affine cost functions where travelers that have heterogeneous value-of-time.

One effort to ameliorate the distributional problems resulting from congestion pricing is to redistribute the toll revenue (see [176, 83, 50, 3, 88, 99], or references therein), or provide tradable or untradable travel credits (see [198, 124, 53, 149, 190], or references therein). The papers [82, 176] were amongst the first to propose different ways to redistribute the revenue in form of infrastructure development and tax rebates. The effectiveness of redistribution schemes are theoretically analyzed in single-lane bottleneck models [10, 24], parallel networks [3], and single origin-destination networks [57].

Pareto-improving congestion pricing schemes were introduced as another approach to reduce inequality. First proposed by [111], Pareto-improving congestion pricing minimizes the total congestion while ensuring that no travelers are worse off in comparison to no tolls. The paper [179] studied the design of Pareto-improving schemes for travelers with heterogeneous value-of-time, and [113] further proved that such Pareto-improving schemes only exist in special classes of networks. The paper [88] studied the problem of designing Pareto-improving pricing schemes combined with a revenue refund. The work [99] extended this line of research by developing optimal revenue refunding schemes to minimize the congestion and distributional problems together. In both [88] and [99], the tolls minimize the weighted sum of travel times with the weights being each population's value-of-time. This objective is different from our goal of minimizing the unweighted total travel time, which is a more suitable metric to assess the environmental impact of congestion on traffic networks.

A third line of works introduced the study of *fairness constrained traffic assignment problems*, proposed by [97], where the fairness metric is the maximum difference of travel time experienced by travelers between the same origin-destination pair. [8, 9] extended this line of research by developing algorithmic methods to solve the fairness constrained traffic assignment problem. The problem of devising congestion pricing schemes which could enforce the resulting traffic assignment patterns was studied in [100]. [100] studies a homogeneous pricing scheme that implements the traffic assignment minimizing an interpolation of the potential function (which is used to characterize the equilibrium) and the social cost function.

This chapter contributes to the above from three aspects: *(i)* Our distributional consideration accounts for both the travel time cost and the monetary cost, which includes both the toll and the gas prices. This generalizes the fairness notion that focuses only on the travel time difference; *(ii)* Our tolling scheme minimizes the total congestion in the network (i.e. guarantees the optimal efficiency) while providing a planner a flexible way to trade-off between the total welfare, the cost distribution across heterogeneous populations and total revenue. In particular, by tuning the parameter that governs the trade-off between the average welfare and the distributional term, we can increase or reduce the revenue collected by the pricing scheme; *(iii)* We provide a comprehensive evaluation of different congestion pricing schemes in terms of efficiency, distribution and revenue using real-world data collected in the San Francisco Bay Area.

Another line of research related to this paper is on developing inverse optimization-based tools to estimate model parameters (such as demand and latency functions) in non-atomic routing games [189, 25, 195]. There are several differences between our approach and these works. First, we use high fidelity datasets to directly estimate the latency on every edge and the demand of travelers. Second, we consider heterogeneous population of travelers as opposed to the homogeneous population of travelers considered in these works.

Finally, on the empirical side, [72, 146, 18, 197] focused on understanding the impact of congestion pricing of the San Francisco-Oakland Bay Bridge, which is the most heavily congested segment in the San Francisco Bay Area highway network. Our work generalizes this line of work to the entire Bay Area highway network using high-fidelity mobility and socioeconomic datasets.

5.3 Model

In this section, we introduce the non-atomic networked routing game model that forms the basis for our theoretical and computational results. We introduce equilibrium routing and the four types of congestion pricing schemes we consider in this paper.

Network

Consider a transportation network $\tilde{\mathcal{G}} = (N, E)$, where N is the set of nodes, and E is the set of edges. A set of non-atomic travelers (agents) make routing decisions in the network between their origin and destination. We denote the set of origin-destination (o-d) pairs by K and the set of routes (i.e. sequences of edges) connecting each o-d pair $k \in K$ by R^k.

Travelers for each o-d pair k are grouped into I populations, where each population is associated with a different level of *value-of-time* $\theta^i \in \mathbb{R}_{\geq 0}$ that captures the trade-off travelers in population i are willing to make between travel time and monetary cost while selecting between different routes. We refer to agents with value-of-time θ^i as type i agents. The demand vector is given by $D = (D^{ik})_{i \in I, k \in K}$, where D^{ik} is the demand of agents with type i that want to travel between o-d pair k. Throughout, we operate under the *inelastic demand* assumption: traveler demands on each origin-destination pair are constant. This assumption is reasonable in the empirical study that follows, given that *(i)* our analysis focuses on the commuting behavior during the morning rush hour, when the majority of trips are work-related with little elasticity; *(ii)* the availability of public transit is sparse and the cost of car ownership is high [52].

The strategy distribution of agents is denoted $q = (q_r^{ik})_{r \in R^k, i \in I, k \in K}$, where q_r^{ik} is the flow of agents with type i and o-d pair k who take route r. Therefore, given a demand D, the set of feasible strategy distributions is given by:

$$\mathcal{Q}(D) := \left\{ q : \sum_{r \in R^k} q_r^{ik} = D^{ik}, q_r^{ik} \geq 0, \forall r \in R^k, i \in I, k \in K \right\}. \tag{5.1}$$

Given a strategy distribution $q \in \mathcal{Q}(D)$, the flow of agents of type $i \in I$ on edge $e \in E$ is given by

$$f_e^i(q) := \sum_{k \in K} \sum_{r \in R^k} q_r^{ik} \mathbb{1}(e \in r), \tag{5.2}$$

and the total flow of agents on edge $e \in E$ is

$$w_e(q) := \sum_{i \in I} f_e^i(q). \tag{5.3}$$

The travel time experienced by agents taking edge $e \in E$ is $\ell_e(w_e(q))$, where the latency function $\ell_e \colon \mathbb{R}_+ \to \mathbb{R}_+$ is *continuous, strictly increasing, and convex.* Consequently, the total travel time experienced by agents from o-d pair $k \in K$ who use route $r \in R^k$ is given by $\ell_r(q) := \sum_{e \in r} \ell_e(w_e(q))$. With slight abuse of notation, we use $\ell_r(q)$ and $\ell_r(w)$ interchangeably to represent the latency of route r where w is the edge flow vector corresponding to the strategy distribution q. In addition to the travel time, the total cost experienced by each individual agent also includes the congestion price imposed by the planner, and the gas cost required to travel on the route the agent chooses. In particular, let p_e^i be the toll price imposed on travelers of type $i \in I$ for using edge $e \in E$, and g_e be the gas cost of using an edge $e \in E$. Note that we allow for the toll price to be type-specific in the general setting. We will later discuss different scenarios for setting the toll prices. Given the tolls $p = (p_e^i)_{e \in E, i \in I}$, the cost experienced by travelers of type $i \in I$ associate with o-d pair $k \in K$ and taking route $r \in R^k$ is given by

$$c_r^i(q, p) := \ell_r(q) + \frac{1}{\theta^i} \sum_{e \in r} (p_e^i + g_e). \tag{5.4}$$

Crucially, a key feature of our model is that the toll and gas costs experienced by each agent are modulated by the value-of-time θ^i of that agent. This allows us to model the heterogeneity present in the types of travelers. Given this setup, we define Nash equilibrium to be the strategy distribution such that no traveler has incentive to deviate from their chosen route. That is,

Definition 5.3.1. *Given tolls p, a strategy profile $q^*(p)$ is a Nash equilibrium if*

$$\begin{aligned} &\forall i \in I, k \in K, r \in R^k, \quad q_r^{ik*}(p) > 0 \\ \Rightarrow \quad & c_r^i(q^*(p), p) \le c_{r'}^i(q^*(p), p) \quad \forall r' \in R^k. \end{aligned}$$

The objective of the planner is to minimize the network congestion, measured by the total travel time experienced by all travelers. For any strategy distribution q, we denote the planner's cost function as follows:

$$S(q) := \quad \sum_{e \in E} w_e(q) \ell_e(w_e(q)), \tag{5.5}$$

where $w_e(q)$ is given by (5.3). We denote the set of socially optimal strategy distributions as $q^\dagger := \operatorname{argmin}_{q \in \mathcal{Q}(D)} S(q)$, and the induced socially optimal edge flows as $w^\dagger = (w_e^\dagger)_{e \in E}$, where $w_e^\dagger = w_e(q^\dagger)$ is given by (5.3).

Congestion pricing

We now introduce two practical considerations for implementing the congestion price. The first consideration is whether or not the toll is type-specific. In particular, a congestion pricing scheme is *homogeneous* if the toll is uniform across all population types, and *heterogeneous* if the toll varies with population types (formally, whether p_e^i is allowed to depend on i or not, on each edge). The challenge of implementing a heterogeneous scheme is that the population type (i.e. value-of-time) is a latent variable that is privately known only by the individual traveler. In practice, an individual's value-of-time sis often closely correlated with their income level, i.e. higher-income groups are typically associated with a higher value-of-time, while lower-income groups correlate with a lower value-of-time [13, 87, 188, 184]. Therefore, one way to implement heterogeneous tolling is to set tolls based on the income level of travelers. For example, low income groups, which have significant overlap with the population of low value-of-time travelers, may receive a subsidy or a toll rebate in certain areas. Such rebate programs have been established in several states in the United States, e.g., California [1] , Virginia [2] , New York [3].

The second consideration is whether or not tolls can be set on all the edges of the network or only on a subset (formally, whether or not p_e^i is allowed to be strictly positive on all $e \in E$). In practical terms, congestion pricing often requires the installation of physical facilities, which might not be feasible on all road segments. Thus, a congestion pricing scheme has no support constraints if tolls can be imposed on all edges, or has support constraints if tolls can only be imposed on a subset of edges, denoted as E_T. We note that congestion pricing schemes with (respectively without) support constraints are also referred to as first-best (respectively second-best) tolling schemes in the literature.

Building on the above two considerations, we define four types of tolling schemes: *(i) Homogeneous tolls with no support constraints* (hom): $p_e^i \geq 0$ and $p_e^i = p_e^j$ for all $e \in E$ and all $i, j \in I$; *(ii) Heterogeneous tolls with no support constraints* (het): $p_e^i \geq 0$ for all $e \in E, i \in I$; *(iii) Homogeneous tolls with support constraints* (hom_sc): $p_e^i = p_e^j$ for all $i, j \in I$ and all $e \in E$. Additionally, $p_e^i = 0$ for all $e \in E \backslash E_T$, and $p_e^i \geq 0$ for all $e \in E_T$; *(iv) Heterogeneous tolls with support constraints* (het_sc): $p_e^i = 0$ for all $e \in E \backslash E_T$, and $p_e^i \geq 0$ for all $e \in E_T$.

5.4 Computation Methods

In this section, we outline methods for computing equilibrium routing strategies and the four congestion pricing schemes. We first establish that, given any fixed toll values, the equilibrium outcome can be derived as the optimal solution to a convex optimization problem. We then demonstrate that the set of homogeneous tolls (hom) and heterogeneous tolls (het) without support constraints that realize the socially optimal edge flows can be characterized as the set of optimal solutions of linear programs. Next, we present a multi-step approach for calculating the toll prices that strikes a balance between the cost distribution, as measured by the cost disparity between travelers from different populations, and at the same time, maximizing the welfare of all traveler populations. For congestion pricing schemes with support constraints, we adapt our approach to provide a heuristic for calculating hom_sc and het_sc, acknowledging that such solutions may not guarantee the implementation of the socially optimal edge flows.

[1] `https://mtc.ca.gov/news/new-year-brings-new-toll-payment-assistance-programs`
[2] `https://www.vdottollrelief.com/`
[3] `https://new.mta.info/fares-and-tolls/bridges-and-tunnels/resident-programs`

Proposition 5.4.1. *Given toll prices p, a strategy distribution $q^*(p)$ is a Nash equilibrium if and only if it is a solution to the following convex optimization problem:*

$$\begin{aligned}\min_{q\in\mathcal{Q}(D)} \quad \Phi(q,p,\theta) = & \sum_{e\in E}\int_0^{w_e(q)} \ell_e(z)\ \mathrm{d}z \\ & + \sum_{i\in I}\sum_{e\in E}\frac{(p_e^i+g_e)}{\theta^i} f_e^i(q),\end{aligned} \tag{5.6}$$

where $w_e^i(q)$, $w_e(q)$ are given by (5.2) *and* (5.3), *respectively. Moreover, given any toll price vector p, the equilibrium edge flow vector $w^*(p) := w(q^*(p))$ is unique. Additionally, the socially optimal edge flow vector $w^\dagger$ is unique.*

[88] showed the same result as Proposition 5.4.1 without the gas price. The proof follows directly from the method in [88].

Proposition 5.4.1 shows that $w^\dagger$ is unique. However, we note that such a $w^\dagger$ may be induced by multiple type-specific flow vectors $f^\dagger$. Although these different type-specific flow vectors all induce the same aggregate edge load, and thus minimize the total cost, they may lead to different travel times experienced by different populations.

The following proposition shows that the set of prices hom (respectively het) that implements the socially optimal edge load can be characterized each by a linear program.

Proposition 5.4.2. *(1) A homogeneous congestion pricing scheme $p^\dagger = (p_e^\dagger)_{e\in E}$ implements the socially optimal edge flow $w^\dagger = (w_e^\dagger)_{e\in E}$ if and only if there exists $z^\dagger$ such that $(p^\dagger, z^\dagger)$ is a solution to the following linear program:*

$$\begin{aligned} T^*_{\textsf{hom}} = \max_{p,z} \ & \sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E} p_e w_e^\dagger, \\ s.t. \quad & z^{ik} - \sum_{e\in r}(p_e+g_e) \le \theta^i \ell_r(w^\dagger), \\ & \forall k\in K, r\in R^k, i\in I, \\ & p_e \ge 0, \quad \forall e\in E. \end{aligned} \tag{$\mathcal{P}_{\textsf{hom}}$}$$

(2) A heterogeneous congestion pricing scheme $p^\dagger = (p_e^{i\dagger})_{e\in E, i\in I}$ implements a type-specific socially optimal edge flow $f^{\dagger,} = (f_e^{\dagger i,})_{e\in E, i\in I}$ if and only if there exists a $z^\dagger$ such that $(p^\dagger, z^\dagger)$ is a solution to the following linear program:

$$\begin{aligned} T^*_{\textsf{het}}(f^\dagger) = \max_{p,z} \ & \sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E}\sum_{i\in I} p_e^i f_e^{\dagger i,}, \\ s.t.\ & z^{ik} - \sum_{e\in r}(p_e^i+g_e) \le \theta^i \ell_r(w^\dagger), \\ & \forall k\in K, r\in R^k, i\in I, \\ & p_e^i \ge 0, \quad \forall e\in E, i\in I. \end{aligned} \tag{$\mathcal{P}_{\textsf{het}}$}$$

Proposition 5.4.2 follows the results in [70, 103, 137, 191, 91]. The proof builds on the two linear programs ($\mathcal{P}_{\mathsf{hom}}$) – ($\mathcal{P}_{\mathsf{het}}$) and their dual programs ($\mathcal{D}_{\mathsf{hom}}$) and ($\mathcal{D}_{\mathsf{het}}$) as follows:

$$
\begin{aligned}
\min_{q} \quad & \sum_{i\in I}\sum_{k\in K}\sum_{r\in R^k}(\theta^i \ell_r(w^\dagger) + \sum_{e\in r} g_e) q_r^{ik} && (\mathcal{D}_{\mathsf{hom}})\\
\text{s.t.} \quad & \sum_{i\in I}\sum_{k\in K}\sum_{r\in R^k: e\in r} q_r^{ik} \le w_e^\dagger, \quad \forall e \in E, && (\mathcal{D}_{\mathsf{hom}.a})\\
& \sum_{r\in R^k} q_r^{ik} = D^{ik}, \quad \forall i \in I, k \in K, && (\mathcal{D}_{\mathsf{hom}.b})\\
& q_r^{ik} \ge 0 \quad \forall i \in I, k\in K, r \in R^k. && (\mathcal{D}_{\mathsf{hom}.c})\\
\min_{q} \quad & \sum_{i\in I}\sum_{k\in K}\sum_{r\in R^k}(\theta^i \ell_r(w^\dagger) + \sum_{e\in r} g_e) q_r^{ik} && (\mathcal{D}_{\mathsf{het}})\\
\text{s.t.} \quad & \sum_{k\in K}\sum_{r\in R^k: e\in r} q_r^{ik} \le f_e^{\dagger i}, \quad \forall e \in E, i \in I, && (\mathcal{D}_{\mathsf{het}.a})\\
& \sum_{r\in R^k} q_r^{ik} = D^{ik}, \quad \forall i \in I, k \in K, && (\mathcal{D}_{\mathsf{het}.b})\\
& q_r^{ik} \ge 0, \quad \forall i \in I, k\in K, r \in R^k. && (\mathcal{D}_{\mathsf{het}.c})
\end{aligned}
$$

Under both hom and het, the feasibility constraints of the associated primal and dual programs as well as the complementary slackness conditions are equivalent to the equilibrium condition where only routes with the minimum cost are taken by travelers. Moreover, constraints ($\mathcal{D}_{\mathsf{hom}.a}$) and ($\mathcal{D}_{\mathsf{het}.a}$) must be tight at optimality, indicating that the induced flow vector in equilibrium is indeed $w^\dagger$, which minimizes total travel time. Therefore, the set of optimal solutions of ($\mathcal{P}_{\mathsf{hom}}$) and ($\mathcal{P}_{\mathsf{het}}$) are the set of toll vectors that induce $w^\dagger$ under hom and het, respectively.

We denote $P^\dagger_{\text{hom}}$ as the set of socially optimal toll prices for hom, and $P^\dagger_{\mathsf{het}}(f^\dagger)$ as the set of socially optimal toll price for het that induces a type-specific socially optimal edge flow $f^\dagger$. Proposition 5.4.2 demonstrates that both sets can be computed as the optimal solution set of linear programs. We note that the set $P^\dagger_{\mathsf{het}}(f^\dagger)$ depends on which type-specific socially optimal flow $f^\dagger$ is induced since the objective function ($\mathcal{P}_{\mathsf{het}}$) depends on $f^\dagger$.

Furthermore, $P^\dagger_{\text{hom}}$ and $P^\dagger_{\mathsf{het}}(f^\dagger)$ may not be singleton. This presents an opportunity for the planner to decide which specific toll price from the optimal solution set to implement. While all tolls in $P^\dagger_{\text{hom}}$ and $P^\dagger_{\mathsf{het}}(f^\dagger)$ achieve the minimum social cost, they do so by impacting travelers differently given their individual origin-destination pair and value-of-time. We consider that the central planner aims at solving the following problem:

$$\begin{aligned}
\min_{p} \ & L(p) := \\
& \underbrace{\max_{i,i' \in I} \left| \frac{1}{D^i} \sum_{k \in K} D^{ik} \frac{c^{ik\dagger}(p)}{c^{ik\dagger}(0)} - \frac{1}{D^{i'}} \sum_{k' \in K} D^{i'k'} \frac{c^{i'k'\dagger}(p)}{c^{i'k'\dagger}(0)} \right|}_{(i)} \\
& \qquad + \lambda \underbrace{\frac{1}{D} \sum_{i \in I} \sum_{k \in K} D^{ik} \frac{c^{ik\dagger}(p)}{c^{ik\dagger}(0)}}_{(ii)}, \\
s.t. \quad & p \in \begin{cases} P^{\dagger}_{\text{hom}}, & \text{in } \mathsf{hom}, \\ P^{\dagger}_{\mathsf{het}}(f^{\dagger}), & \text{in } \mathsf{het} \text{ with route flow } f^{\dagger}, \end{cases}
\end{aligned} \tag{5.8}$$

where $D^i = \sum_{k \in K} D^{ik}$, $D = \sum_{i \in I} D^i$, $\lambda \geq 0$,

$$c^{ik\dagger}(p) = \min_{r \in R^k} \left\{ \ell_r(w^{\dagger}) + \frac{1}{\theta^i} \sum_{e \in r} (p_e + g_e) \right\} \tag{5.9}$$

is the equilibrium cost of individuals with o-d pair k and type i given the toll price p and socially optimal edge load $w^{\dagger}$, and $c^{ik\dagger}(0)$ is the equilibrium cost of individuals with o-d pair k and type i given no (or zero) tolls. We emphasize that the cost $c^{ik\dagger}(p)$ is the minimum cost of choosing a route given the socially optimal load vector $w^{\dagger}$, the toll price, and the gas fee. This is indeed an equilibrium cost of traveler population with type i and o-d pair k since any $p \in P^{\dagger}_{\text{hom}}$ or $p \in P^{\dagger}_{\mathsf{het}}(f^{\dagger})$ guarantees that the equilibrium edge vector is $w^{\dagger}$.

The objective function in (5.8) indicates that the central planner selects the toll price that not only minimizes the total travel time but also balances the distribution of costs among populations with different values-of-time, and the average welfare that accounts for the travel time as well as the toll price and gas fee. In particular, in (5.8) *(i)* reflects an objective that assesses the maximum disparity in the relative change in travel costs experienced by different types of travelers following the implementation of tolls, compared to a scenario without tolls, and *(ii)* reflects an average welfare objective that is the average of the relative change in travel costs experienced by all of the travelers following the implementation of tolls, compared to a scenario without tolls. Balancing welfare maximization with cost disparity minimization avoids the potential problem with just minimizing cost disparity: charging excessively high prices to all types of travelers. Moreover, $\lambda \geq 0$ is a parameter that governs the relative weight between the distributional objective and the welfare objective.

We denote the socially optimal homogeneous congestion pricing scheme that solves the central planner's problem (5.8) as p^{*}_{hom}. The next proposition shows that we can solve the central planner's problem (5.8) for hom by another linear program.

Proposition 5.4.3. *For the* hom *tolling scheme, p^*_{hom} is an optimal solution of the following linear program:*

$$\min_{p,z,y} \quad y + \frac{\lambda}{D}\sum_{i\in I}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} \qquad (\mathcal{P}^*_{\mathsf{hom}})$$

$$\text{s.t.} \quad y \geq \frac{1}{D^i}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} - \frac{1}{D^{i'}}\sum_{k'\in K} D^{i'k'}\frac{z^{i'k'}}{\theta^{i'} c^{i'k'\dagger}(0)}, \quad \forall i, i' \in I, \qquad (\mathcal{P}^*_{\mathsf{hom}.a})$$

$$\sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E} p_e w^\dagger_e \geq T^*_{\mathit{hom}}, \qquad (\mathcal{P}^*_{\mathsf{hom}.b})$$

$$z^{ik} - \sum_{e\in r}(p_e + g_e) \leq \theta^i \ell_r(w^\dagger), \ \forall k \in K, r \in R^k, i \in I, \qquad (\mathcal{P}^*_{\mathsf{hom}.c})$$

$$p_e \geq 0 \quad \forall e \in E, \qquad (\mathcal{P}^*_{\mathsf{hom}.d})$$

where T^*_{hom} is the optimal value of $(\mathcal{P}_{\mathsf{hom}})$.

In $(\mathcal{P}^*_{\mathsf{hom}})$, constraints $(\mathcal{P}^*_{\mathsf{hom}.c})$ and $(\mathcal{P}^*_{\mathsf{hom}.d})$ ensure that variables (p, z) are in the feasible set of $(\mathcal{P}_{\mathsf{hom}})$, and constraint $(\mathcal{P}^*_{\mathsf{hom}.b})$ further restricts the set of (p, z) in $(\mathcal{P}^*_{\mathsf{hom}})$ to be the set of optimal solutions of $(\mathcal{P}_{\mathsf{hom}})$. Thus, following Proposition 5.4.2, any feasible p in $(\mathcal{P}^*_{\mathsf{hom}})$ must be a toll vector that induces the socially optimal edge flow $w^\dagger$. Moreover, the proof of Proposition 5.4.2 further ensures that for every $i \in I, k \in K$ there exists $r \in R^k$ such that the corresponding constraint in $(\mathcal{P}^*_{\mathsf{hom}.c})$ must be tight at optimality, which indicates that any z^{ik} in $(\mathcal{P}^*_{\mathsf{hom}})$ equals to $\theta^i \cdot c^{ik\dagger}(p)$. Additionally, constraints $(\mathcal{P}^*_{\mathsf{hom}.a})$ guarantee that at optimality $y = \max_{i,i'\in I}\left|\frac{1}{D^i}\sum_{k\in K} D^{ik}\frac{c^{ik\dagger}(p)}{c^{ik\dagger}(0)} - \frac{1}{D^{i'}}\sum_{k'\in K} D^{i'k'}\frac{c^{i'k'\dagger}(p)}{c^{i'k'\dagger}(0)}\right|$. Thus, the linear program $(\mathcal{P}^*_{\mathsf{hom}})$ computes the homogeneous toll prices that minimize the total travel time and optimize the distributional and welfare objectives with a relative weight λ.

To summarize, the programs $(\mathcal{P}_{\mathsf{hom}})$ and $(\mathcal{P}^*_{\mathsf{hom}})$ provide a *two-step approach* for computing p^*_{hom}: first, compute T^*_{hom} by solving the linear program $(\mathcal{P}_{\mathsf{hom}})$ given the unique edge flow $w^\dagger$. Second, compute p^*_{hom} by solving the linear program $(\mathcal{P}^*_{\mathsf{hom}})$ using T^*_{hom}.

Next, we show that the planner can compute the heterogeneous toll price vector (het) that minimize the total travel time and optimize the distribution-welfare objective (5.8), denoted by p^*_{het}, using a similar approach as described above. However, in het, one additional issue arises, since the set $P^\dagger_{\mathsf{het}}(f^\dagger)$ and consequently p^*_{het} depend on the selection of the type-specific flow vector $f^\dagger$, which may not be unique. Here, we propose to select the type-specific flow vector $f^\dagger$ as the one that induces the edge flow vector $w^\dagger$ (which minimizes total travel time) while also minimizing the disparity in the total travel time experienced across all traveler populations. To compute such a $f^\dagger$, we first find a feasible routing strategy profile $q^\dagger$ that induces $w^\dagger$ and minimizes the average cost difference among traveler populations. Such a $q^\dagger$

can be solved by the following linear program:

$$\begin{aligned}
\min_{q} \quad & x, \\
\text{s.t.} \quad & x \geq \sum_{k\in K}\sum_{r\in R^k}\left(q_r^{ik}\ell_r(w^\dagger) - q_r^{i'k}\ell_r(w^\dagger)\right), \forall i, i' \in I, \\
& \sum_{r\in R^k} q_r^{ik} = D^{ik}, \forall\ i \in I, k \in K, \\
& \sum_{i\in I}\sum_{k\in K}\sum_{r\in R^k: e\in r} q_r^{ik} = w_e^\dagger, \forall e \in E, \\
& q_r^{ik} \geq 0, \forall\ i \in I, k \in K, r \in R^k.
\end{aligned}$$

Then, the induced population-specific flow vector $f^\dagger$ associated with $q^\dagger$ is given by (5.2). Based on $f^\dagger$, we compute p^*_{het} as the optimal solution of a linear program.

Proposition 5.4.4. *For the* het *tolling scheme, given $f^\dagger$, p^*_{het} is an optimal solution of the following linear program:*

$$\min_{p,z,y} \quad y + \frac{\lambda}{D}\sum_{i\in I}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} \qquad (\mathcal{P}^*_{\mathsf{het}})$$

$$s.t. \quad y \geq \frac{1}{D^i}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} - \frac{1}{D^{i'}}\sum_{k'\in K} D^{i'k'}\frac{z^{i'k'}}{\theta^{i'} c^{i'k'\dagger}(0)}, \quad \forall i, i' \in I, \qquad (\mathcal{P}^*_{\mathsf{het}.a})$$

$$\sum_{i\in I}\sum_{k\in K} D^{ik}z^{ik} - \sum_{e\in E} p_e w_e^\dagger \geq T^*_{\mathsf{het}}(f^\dagger), \qquad (\mathcal{P}^*_{\mathsf{het}.b})$$

$$z^{ik} - \sum_{e\in r}(p_e^i + g_e) \leq \theta^i \ell_r(w^\dagger), \forall k \in K, r \in R^k, i \in I, \qquad (\mathcal{P}^*_{\mathsf{het}.c})$$

$$p_e^i \geq 0, \forall e \in E, i \in I, \qquad (\mathcal{P}^*_{\mathsf{het}.d})$$

where $T^*_{\mathsf{het}}(f^\dagger)$ is the optimal value of the objective function of $(\mathcal{P}_{\mathsf{het}})$ associated with $f^\dagger$.

Propositions 5.4.2 and 5.4.4 show that p^*_{het} can be computed using a *three-step approach*: first, we compute the type-specific flow vector $f^\dagger$ that induces the edge flow $w^\dagger$ while also minimizing the travel time difference among all traveler populations using (5.4). Second, we compute $T^*_{\mathsf{het}}(f^\dagger)$ using $(\mathcal{P}_{\mathsf{het}})$ given $f^\dagger$. Third, we compute p^*_{het} using $(\mathcal{P}^*_{\mathsf{het}})$.

Finally, we discuss how to extend our approaches of computing p^*_{hom} and p^*_{het} to incorporate the support constraints of the toll price. Previous studies [95, 29] showed that the problem of computing toll prices that satisfy support set constraints and also minimize the total travel time is NP hard even without considering heterogeneous values-of-time of travelers or distributional objectives. Here, we provide heuristics for computing the toll prices with support constraints. We evaluate the performance of our heuristics in terms of total travel time, cost distribution, and welfare on the San Francisco Bay Area network in Sec. 5.5.

Heuristics for computing $p^*_{\mathsf{hom\_sc}}$ We propose a two-step heuristic to compute $\mathsf{hom\_sc}$ by appropriately modifying the two-step method to compute hom.

We first solve the following linear program that adds the support constraints to $(\mathcal{P}_{\mathsf{hom}})$:

$$
\begin{aligned}
T^*_{\mathsf{hom\_sc}} = \max_{p,z} \ & \sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E} p_e w_e^\dagger, \\
\text{s.t.} \quad & z^{ik} - \sum_{e\in r}(p_e + g_e) \le \theta^i \ell_r(w^\dagger), \\
& \forall k \in K, r\in R^k, i\in I, \\
& p_e \ge 0, \quad \forall e \in E_T, \quad p_e = 0, \quad \forall e\in E\setminus E_T.
\end{aligned} \tag{$\mathcal{P}_{\mathsf{hom\_sc}}$}
$$

We note that the equilibrium edge load associated with any optimal solution of $(\mathcal{P}_{\mathsf{hom\_sc}})$, say $\hat{w}$, may not be equal to the socially optimal edge load $w^\dagger$. This is because the constraints that require edges in $E\setminus E_T$ have zero tolls remove the dual constraints in $(\mathcal{D}_{\mathsf{hom}.a})$ for edges in E_T. As a result, the primal and dual argument in the proof of Proposition 5.4.2 no longer holds, and thus the induced edge flow $\hat{w}$ may not be equal to $w^\dagger$.

Note that the optimal solution to $(\mathcal{P}_{\mathsf{hom\_sc}})$ will be non-unique. Therefore, inspired by $(\mathcal{P}^*_{\mathsf{hom}})$, we consider the following heuristic to incorporate both the distribution and welfare metric while also accounting for support constraints. Note that simply adding the support constraints in $(\mathcal{P}^*_{\mathsf{hom}})$ could render the optimization problem infeasible as the optimal set of homogeneous tolls P^*_{hom} need not have a solution that satisfies the support constraints. In particular, the constraint $(\mathcal{P}^*_{\mathsf{hom}.b})$ would get violated. This is because $T^*_{\mathsf{hom}} \ge T^*_{\mathsf{hom\_sc}}$ as the constraint set of $(\mathcal{P}^*_{\mathsf{hom\_sc}})$ is contained in that of $(\mathcal{P}^*_{\mathsf{hom}})$. Therefore, we compute $p^*_{\mathsf{hom\_sc}}$ as the optimal solution of the following linear program which adds support constraints to $(\mathcal{P}^*_{\mathsf{hom}})$ while relaxing the constraint $(\mathcal{P}^*_{\mathsf{hom}.b})$ by using $T^*_{\mathsf{hom\_sc}}$ instead of T^*_{hom}:

$$
\min_{p,z,y} \quad y + \frac{\lambda}{D}\sum_{i\in I}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} \tag{$\mathcal{P}^*_{\mathsf{hom\_sc}}$}
$$

$$
\text{s.t.} \quad y \ge \frac{1}{D^i}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} - \frac{1}{D^{i'}}\sum_{k'\in K} D^{i'k'}\frac{z^{i'k'}}{\theta^{i'} c^{i'k'\dagger}(0)}, \quad \forall i, i' \in I, \tag{$\mathcal{P}^*_{\mathsf{hom\_sc}.a}$}
$$

$$
\sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E} p_e w_e^\dagger \ge T^*_{\mathsf{hom\_sc}}, \tag{$\mathcal{P}^*_{\mathsf{hom\_sc}.b}$}
$$

$$
z^{ik} - \sum_{e\in r}(p_e+g_e) \le \theta^i \ell_r(w^\dagger), \forall k\in K, r\in R^k, i\in I, \tag{$\mathcal{P}^*_{\mathsf{hom\_sc}.c}$}
$$

$$
p_e \ge 0, \quad \forall e\in E_T, p_e = 0, \quad \forall e \in E\setminus E_T. \tag{$\mathcal{P}^*_{\mathsf{hom\_sc}.d}$}
$$

Heuristics for computing $p^*_{\mathsf{het\_sc}}$ The computation of $p^*_{\mathsf{het\_sc}}$ follows a three-step procedure, similar to that of p^*_{het} but restricting the set of allowable tolls to be zero on non-tollable edges, as done in hom_sc. First, we compute the population-specific flow vector $f^\dagger$ that induces the congestion minimizing edge flow $w^\dagger$ while also minimizes the average difference of travel time among all traveler populations using (5.4). Next, we add the support constraints to $(\mathcal{P}_{\textsf{het}})$ to compute the optimal value $T^*_{\textsf{het\_sc}}(f^\dagger)$ as follows:

$$\begin{aligned} T^*_{\textsf{het\_sc}}(f^\dagger) = \max_{p,z} \ & \sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E}\sum_{i\in I} p^i_e f^{\dagger i}_e, \\ \text{s.t. } & z^{ik} - \sum_{e\in r}(p^i_e + g_e) \leq \theta^i \ell_r(w^\dagger), \\ & \quad \forall k \in K, r \in R^k, i \in I, \\ & p^i_e \geq 0, \quad \forall e \in E_T \ \forall i \in I, \\ & p_e = 0, \quad \forall e \in E \setminus E_T \ \forall i \in I. \end{aligned} \tag{$\mathcal{P}_{\textsf{het\_sc}}$}$$

Analogous to the case hom_sc, the equilibrium edge load associated with the optimal solution of $(\mathcal{P}_{\textsf{het\_sc}})$, $\hat{w}$, may not be equal to the socially optimal edge load $w^\dagger$ due to the added support constraints. We compute $p^*_{\textsf{het\_sc}}$ as the optimal solution of the following linear program which adds support constraints to $(\mathcal{P}^*_{\textsf{het}})$ while relaxing the constraint $(\mathcal{P}^*_{\textsf{het}.b})$ by using $T^*_{\textsf{het\_sc}}$ instead of $T^*_{\textsf{het}}$:

$$\min_{p,z,y} \quad y + \frac{\lambda}{D}\sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik}, \tag{$\mathcal{P}^*_{\textsf{het\_sc}}$}$$

$$\text{s.t.} \quad y \geq \frac{1}{D^i}\sum_{k\in K} D^{ik}\frac{z^{ik}}{\theta^i c^{ik\dagger}(0)} - \frac{1}{D^{i'}}\sum_{k'\in K} D^{i'k'}\frac{z^{i'k'}}{\theta^{i'} c^{i'k'\dagger}(0)}, \quad \forall i, i' \in I, \tag{$\mathcal{P}^*_{\textsf{het\_sc}.a}$}$$

$$\sum_{i\in I}\sum_{k\in K} D^{ik} z^{ik} - \sum_{e\in E} p^i_e w^\dagger_e \geq T^*_{\textsf{het\_sc}}, \tag{$\mathcal{P}^*_{\textsf{het\_sc}.b}$}$$

$$z^{ik} - \sum_{e\in r}(p^i_e + g_e) \leq \theta^i \ell_r(w^\dagger), \quad \forall k \in K, r \in R^k, i \in I, \tag{$\mathcal{P}^*_{\textsf{het\_sc}.c}$}$$

$$p^i_e \geq 0, \ \forall e \in E_T, i \in I, \quad p^i_e = 0, \ \forall e \in E \setminus E_T, i \in I. \tag{$\mathcal{P}^*_{\textsf{het\_sc}.d}$}$$

5.5 Model Calibration for the San Francisco Bay Area

In this section, we calibrate the non-atomic routing game model for the San Francisco Bay Area freeway network using the Caltrans Performance Measurement System (PeMS) dataset [4], American Community Survey (ACS) dataset [5] and Safegraph neighborhood patterns dataset from 2019 [6]. We briefly describe each dataset. We subsequently present the

[4] Available at `https://pems.dot.ca.gov/`

[5] Available at `https://www.census.gov/programs-surveys/acs`

[6] This dataset was available for public use at `https://www.safegraph.com` till 2021 and is now commercially available

calibration of the Bay Area transportation network, the demand of each population type, and the value-of-time parameters.

Datasets

Caltrans PeMS Dataset The Caltrans PeMS dataset is based on measurements taken from loop detectors placed on a network of freeways and bridges in California. Our dataset is taken from district 4, which covers the entire San Francisco Bay Area. This dataset provides hourly flow counts and average vehicle speeds measured by each loop detector placed along the freeways. We use this dataset to calibrate the latency functions of edges.

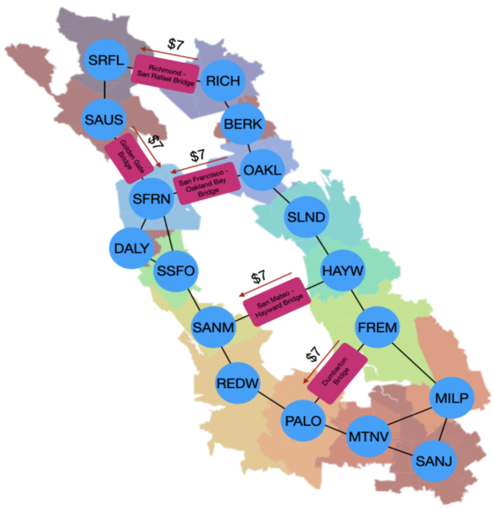
American Community Survey (ACS) Dataset The ACS dataset is collected by the US Census Bureau to record demographic and socioeconomic information. We use the information from the Means of Transportation (2019) entry in ACS, which provides information of commuters' mode choices (percentage of driving population), employment, and household income. The dataset is collected at the zip-code level for the entire United States.

Safegraph Neighborhood Patterns Dataset This dataset records the aggregate mobility pattern using the data collected from 40 million mobile devices in the US. This dataset estimates the commuting pattern by counting the number of mobile devices that travel from one census block group (CBG) to another CBG and dwell for at least 6 hours between 7:30 am and 5:30 pm Monday through Friday. We use this dataset in conjunction with the ACS dataset to estimate the demand of driving commuters between each o-d pair in the network within each income level.

The San Francisco Bay Area freeway network

We represent the San Francisco Bay Area using a network with 17 nodes (see Fig. 5.3). Each node represents a major city, and the edges are the major freeways connecting these cities. Among these edges, five of them are bridges: the Golden Gate Bridge, the Richmond-San Rafael Bridge, the San Francisco-Oakland Bay Bridge, the San Mateo-Hayward Bridge, and the Dumbarton Bridge. They are represented as the magenta boxes in Figure 5.3. In 2019, a flat toll of \$7 is imposed for a single crossing on each bridge in the direction denoted in Figure 5.3.

Demand estimate We categorize the driving population into three distinct segments based on their value-of-time, namely *low, middle, and high* value-of-time. The determination of the fraction of driving population in each of these categories relies on the *Means of Transportation* dataset from ACS. Specifically, we assign a traveler to the (a) low value-of-time category if their annual individual income is less than $\$25,000$, to the (b) middle value-of-time

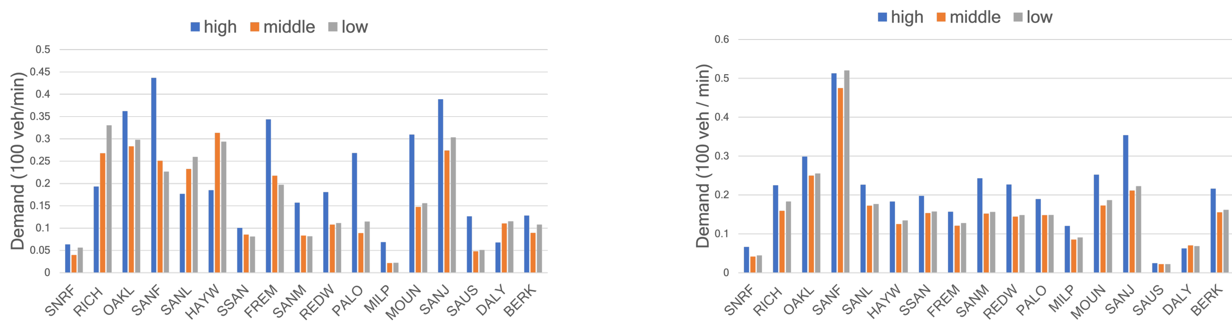


Node	Abbr
San Rafael	SRFL
Richmond	RICH
Oakland	OAKL
San Francisco	SFRN
San Leandro	SLND
Hayward	HAYW
South SF	SSFO
Fremont	FREM
San Mateo	SANM
Redwood City	REDW
Palo Alto	PALO
Milpitas	MILP
Mountain View	MTNV
San Jose	SANJ
Sausalito	SAUS
Daly City	DALY
Berkeley	BERK

Figure 5.3: Bay Area transportation network with tolled bridge segments. Different colors on the map represent the boundaries of cities. The table contains the names of the nodes in map along with abbreviations.

category if their annual individual income falls within the range of \$25,000 to \$65,000, and to the (c) high value-of-time category if their annual individual income exceeds \$65,000.

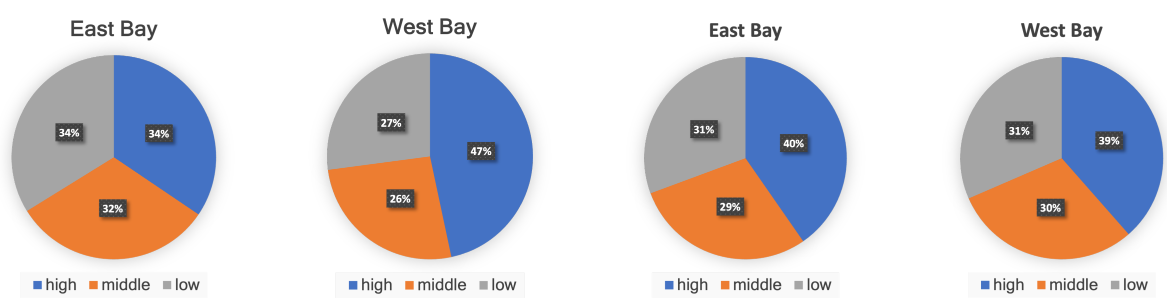
Figure 5.4 provides a visual representation of the distribution of traveler demand to and from each node in the network, stratified by value-of-time. Note that this demand specifically pertains to inter-node travel, with within-node demand excluded from the analysis. We find that approximately 40% of travelers are high willingess-to-pay, and 30% of travelers are of middle and low value-of-time, each. In Figure 5.4a (respectively Figure 5.4b), we present the distribution of traveler demand based on their home (respectively work) location. Around 55% of traffic emerges from relatively few nodes on the East Bay such as RICH, OAKL, SLND, HAYW, FREM, SANJ. Moreover, around 40% of traffic has a work destination in one of the four nodes SFRN, PALO, MTNV, and OAKL. Notably, there exists substantial heterogeneity in both the home and work locations of different traveler types, as can be observed by comparing the distribution of demands in Figure 5.4 to the distribution of median income found in Figure 5.1. For instance, nodes such as RICH, HAYW, SLND, and DALY are predominantly inhabited by a higher number of low value-of-time travelers, while nodes such as PALO, OAKL, SFRN, FREM, and SAUS are predominantly inhabited by high value-of-time travelers. It is interesting to note that on most of the nodes the demographics of



(a) Distribution of traveler types based on their origin (home) nodes.

(b) Distribution of traveler types based on their destination (work) nodes.

Figure 5.4: Distribution of origin and destination traveler demands.



(a) Home Location (b) Home Location (c) Work Location (d) Work Location

Figure 5.5: Distribution of home and work demands across high, middle, and low income levels aggregated over the two sides of the Bay Area.

incoming traffic predominantly comprises high value-of-time travelers. Additionally, as can be seen in Figure 5.5, high-income travelers make up a large fraction demand that originates in the West Bay, as well as of the work location demand on both the East and West Bay.

Next, we describe the approach used to compute the daily demand of different types of travelers traveling between different o-d pairs during January 2019-June 2019. There are three main steps to our approach: first, we obtain an estimate of the relative demand of travelers traveling between different zip-codes in the Bay Area by using the Safegraph dataset. For every month, the Neighborhood Patterns data in the Safegraph dataset provides the average daily count of mobile devices that travel between different census block groups (CBGs) during the work day, which is then aggregated to obtain the relative demand of travelers traveling between different zip codes. After accounting for the sampling bias induced due to the randomly sampled population across the United States, we calibrate demands by using the ACS dataset which provides the income-stratified driving population in every zip code. Finally, to obtain an estimate of daily variability in demand we further augment the demand data with the PeMS dataset by adjusting for daily variation in the total flow on the network in every month. The details of demand estimation are included in [129].

Calibrating the edge latency functions We calibrate the latency functions of each edge of the Bay Area freeway network shown in Figure 5.3. We adopt the Bureau of Public Roads (BPR) function proposed by the Federal Highway Administration (FHA) [136],defined as $\ell_e(w_e) = a_e + b_e w_e^4$, for every $e \in E$, where a_e represents the free-flow travel time (i.e. latency with zero flow) of edge e and b_e is the slope of congestion.

We compute the average driving time of each edge during the morning rush hour (6am to 12pm) on each workday from January 1, 2019 to June 30, 2019 using the speed and distance data from the PeMS dataset. We denote the set of all days as $\mathcal{T}$, the travel time and traffic flow of each edge $e \in E$ on day $t \in \mathcal{T}$ as $\hat{\ell}_e^t$ and $\hat{w}_e^t$, respectively. The details of computing $\left(\hat{\ell}_e^t, \hat{w}_e^t\right)_{t \in \mathcal{T}}$ are provided in [129]. We estimate the free-flow travel time a_e of each $e \in E$ using the average travel time of edge e computed from the PeMS dataset at 3am, when the traffic flow is approaching zero. We denote the estimated value of a_e as $\hat{a}_e$ for each $e \in E$. We next estimate the slope b_e of each edge $e \in E$ using an ordinary least squares regression. In particular, the estimate $\hat{b}_e$ is solved as the minimizer of the following convex program: for every $e \in E$, $\hat{b}_e = \underset{b_e \in \mathbb{R}}{\arg\min} \sum_{t \in \mathcal{T}} \|\hat{\ell}_e^t - \hat{a}_e - b_e \cdot (\hat{w}_e^t)^4\|^2$.

Estimating the value-of-time parameters

We formulate the problem of estimating the value-of-time parameters as an inverse optimization problem. Specifically, the optimal estimate of value-of-time parameters corresponding to the three types of travelers, $\theta^{H*}, \theta^{M*}, \theta^{L*}$, are the ones that minimize the difference between the observed flows on each edge of the network and the corresponding equilibrium edge flows. That is,

$$
\begin{aligned}
\theta_H^*, \theta_M^*, \theta_L^* = \underset{\theta^H, \theta^M, \theta^L}{\operatorname{argmin}} \quad & \sum_{t \in \mathcal{T}} \sum_{e \in E} (\hat{w}_e^t - w_e(q^t))^2 \\
\text{s.t.} \quad & q^t \in \underset{q \in \mathcal{Q}(D^t)}{\arg\min} \ \Phi(q, p, \theta) \quad \forall t \in \mathcal{T}, && (5.14a) \\
& w_e(q^t) \text{ is given by (5.3)}, && (5.14b) \\
& \mathcal{Q}(D^t) \text{ is given by (5.1)}, && (5.14c)
\end{aligned}
$$

where p is the toll price vector in 2019 (i.e. \$7 on each bridge, and \$0 for the remaining edges), $\hat{w}_e^t$ is the observed edge flow on each edge $e \in E$ and each day $t \in \mathcal{T}$ computed using the PeMS dataset, and D^t is the estimated demand vector of each day t computed using the ACS and Safegraph datasets.

Directly solving (5.14) is challenging due to the non-linearity of the edge latency function and the potential function in (5.14a). We compute the estimates using grid search:

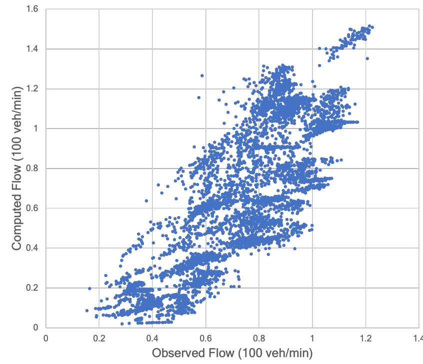


Figure 5.6: Observed and computed equilibrium edge flow.

we construct a grid of value-of-time, where the granularity of each of $\theta^H, \theta^M, \theta^L$ is \$5 per hour. We also assume that the maximum value of value-of-time is \$100 per hour and the minimum is \$0 per hour. Therefore, we define the set of all possible parameter values as $\Theta := \{0, 5, 10, 15, \ldots, 100\}^3$. For each $\theta = (\theta^H, \theta^M, \theta^L) \in \Theta$, we compute the equilibrium flow q_t for every $t \in \mathcal{T}$ and compute the total squared error as in the objective function of (5.14). The optimal parameter θ^* is the one that minimizes the total squared error. We obtain:

$$\theta^* = (\theta^{L*}, \theta^{M*}, \theta^{H*}) = (\$10/\text{hour},\ \$30/\text{hour},\ \$70/\text{hour}).$$

Our estimate θ^* is consistent with the observations reported in prior works, which show that the value-of-timetypically lie between $60\% - 100\%$ of the average hourly income of the population ([154, 13, 140]).

Furthermore, as a robustness check, we plot the equilibrium edge flow $w_e(q^{t*})$ and observed edge flow $\hat{w}_e^t$ for every $e \in E, t \in \mathcal{T}$ in Figure 5.6. Each dot in this figure represents the flow on an edge $e \in E$ on a single day $t \in \mathcal{T}$. Overall, the dots are distributed along the diagonal of the plot indicating that the computed equilibrium edge flows are relatively consistent with the observed edge flows subject to noise in time costs and demand fluctuations.

5.6 Analysis

Our goals in this section are threefold. First, we analyze the average travel time[7] induced at equilibrium due to the current congestion pricing scheme, curr, and identify corridors in the Bay Area which are congested. Next, using the computational method and the calibrated model of the San Francisco Bay area freeway network introduced earlier, we compute the toll values under the congestion pricing schemes hom, het, hom_sc, and het_sc. Finally, we compare different congestion pricing schemes in terms of efficiency and the distribution of travel cost, and also in terms of the overall revenue generated at equilibrium.

[7]We define average travel time to be the ratio between total travel time and the total demand of travelers.

Congestion under the current congestion pricing scheme (curr)

Here, we analyze the average travel time induced at equilibrium under the current congestion pricing scheme, curr, which imposes a uniform toll of $7 on each of the five bridges in the Bay Area, namely on the Richmond-San Rafael Bridge (RICH-SRFL), San Francisco-Oakland Bay Bridge (OAKL-SFRN), Golden Gate Bridge (SAUS-SFRN), San Mateo-Hayward Bridge (HAYW-SANM), and Dumbarton Bridge (FREM-PALO).

Figure 5.7a depicts the difference between the equilibrium travel time given curr and the travel time induce at $w^\dagger$ (normalized by free flow travel time on every edge). We observe that edges on the eastern corridor (connecting nodes RICH-BERK-OAKL-SLND-HAYW-FREM) are over-congested. Meanwhile, the edges on the western corridor (connecting nodes SRFL-SAUS-SFRN-DALY-SSFO-SANM-REDW) are relatively less congested. Furthermore, we observe that amongst all bridges the Bay Bridge (OAKL-SFRN) is also most congested, which is consistent with several prior studies [146, 18, 81]. Additionally, Figure 5.7b presents the difference in the edge flows induced at equilibrium with that of socially optimal edge flows. We observe that in order to reduce the overall congestion we need to ensure that: *(i)* the travelers using the edges in the corridor RICH-BERK-OAKL-SFRN (respectively SFRN-OAKL-BERK-RICH) are incentivized to use the edges in the corridor RICH-SRFL-SAUS-SFRN (respectively SFRN-SAUS-SRFL-RICH), *(ii)* the travelers using the edges in the corridor SFRN-SSFO are incentivized to use the corridor SFRN-DALY-SSFO, and *(iii)* the travelers using the eastern corridor MILP-FREM-HAYW-SLND-OAKL are diverted to use the western corridor MTNV-PALO-REDW-SANM-SSFO by suitably incentivizing them to use the Dumbarton Bridge (FREM-PALO) or the San Mateo-Hayward Bridge(HAYW-SANM).

Furthermore, we note that the average travel cost[8] (the sum of the travel time cost and the equivalent time cost of the monetary expense, as in equation (5.4)) experienced by different types of travelers at equilibrium is unequal in curr. Specifically, low value-of-time travelers bear the travel cost of approximately 91 minutes, while high and middle value-of-time travelers face costs of 61 and 68 minutes, respectively. Moreover, as indicated in Table 5.1, this unequal distribution of travel time persists not only on average but also when examined across different threshold levels of travel cost.

To summarize, we observe that the current congestion pricing scheme implemented the Bay area does not result in efficient allocation of traffic on the network. Additionally, it also leads to unequal distribution of travel cost across different types of travelers.

Toll values under different congestion pricing schemes

Here, using the calibrated model of the Bay Area obtained in Section 5.5, we present the computed values of tolls on various edges of the Bay area network under different congestion pricing schemes (namely, hom, het, hom_sc, het_sc) obtained using the computational method-

[8]It can be shown that the average travel cost experienced by travelers is independent of the route flows on the network and is only dependent on the equilibrium edge flows, which are unique as shown in Proposition 5.4.1.

Travel Cost	Low (%)	Middle (%)	High (%)
$\geq$ 60 minutes	69	55	46
$\geq$ 90 minutes	51	31	28
$\geq$ 120 minutes	32	13	12
$\geq$ 150 minutes	17	1	1

Table 5.1: Fraction of low, middle and high value-of-time travelers that incur total cost (in minutes) more than stated threshold at equilibrium.

ology presented earlier. Figure 5.8a presents the toll values computed under hom by solving ($\mathcal{P}^*_{\mathsf{hom}}$). Figures 5.8b-5.8d present the toll values for low, middle, and high value-of-time travelers under het by solving ($\mathcal{P}^*_{\mathsf{het}}$). Figure 5.8e presents the toll values computed under hom_sc by solving ($\mathcal{P}^*_{\mathsf{hom\_sc}}$). Figure 5.8f further presents the toll values for low, middle, and high value-of-time travelers under het_sc. To compute all of these toll values, we choose $\lambda = 20$ in ($\mathcal{P}^*_{\mathsf{hom}}$), ($\mathcal{P}^*_{\mathsf{het}}$), ($\mathcal{P}^*_{\mathsf{hom\_sc}}$).

Note that in hom and het, on all the bridges, tolls in the east-to-west direction are lower than tolls in the west-to-east direction. This is in contrast to curr, where the west-to-east direction is not tolled at all on any bridge and only the east-to-west direction is tolled at a flat rate of $7 (refer Figure 5.3). Given that the western corridor is less congested than the eastern corridor in curr (refer Figure 5.7a), such tolling is useful to efficiently redistribute traffic in the network. Furthermore, note that in all of the congestion pricing schemes we compute, unlike curr, the Golden Gate Bridge (SAUS-SFRN) is not tolled at all. This choice ensures that more travelers in the eastern corridor, particularly in nodes such as RICH and BERK are able to reach nodes in the west, particularly SFRN, through the Golden Gate Bridge instead of the San Francisco-Oakland Bay Bridge (OAKL-SFRN).

Discussion on efficiency, distribution and revenue generation

In this subsection, we compare the effectiveness of curr, hom, hom_sc, het and het_sc in terms of efficiency (the average travel time for all travelers), equity (average increase in travel cost in comparison to no tolls), and revenue generation (the total toll revenue generated by these schemes). Additionally, we also compare these pricing schemes with the scenario when no toll is implemented (denoted as zero).

Efficiency and Distribution Considerations

Figure 5.9 represents the average travel time experienced by travelers under different congestion pricing schemes. As expected from Proposition 5.4.2, the congestion pricing schemes hom and het achieves the minimum average travel time on the network. Additionally, we note that hom_sc and het_sc do not achieve the minimum average travel time due to the support constraints. Furthermore, it's noteworthy that het_sc results in a slightly improved average travel time compared to hom_sc. This improvement can be attributed to the flexibility of heterogeneous pricing schemes, which allow for type-specific tolls.

Furthermore, from Figure 5.9, we observe that the Price of Anarchy (PoA) – which is the ratio of the social cost of equilibrium value of congestion induced under no tolls with that of the optimum– is 1.04 for the Bay Area transportation network. This is likely due to the high congestion level of the network during the morning rush hour. Theoretical studies (e.g. [42]) have proved that the PoA approaches 1 as the total demand of travelers increases. Moreover, empirical studies (e.g. [150, 143]) have also shown that the PoA in the transportation networks of London, Boston, New York City and Singapore are also close to 1. Furthermore, from Figure 5.9, we find that all congestion pricing schemes hom, hom_sc, het, het_sc outperform curr in terms of average travel time. Surprisingly, curr is also marginally outperformed by zero. A key reason is that curr imposes the same tolls on all of the bridges which does not result in effective re-distribution of traffic from the eastern corridor to the western corridor. While a reduced toll price or zero toll price may increase the total demand of travelers, its impact is likely to be not significant due to (1) the high expense of car ownership and parking fee ([52] estimates that US average annual car ownership cost is $12182 in 2023), and (2) the low coverage of public transportation in the Bay Area.

Figure 5.10 illustrates the average travel cost experienced by type of travelers under different pricing schemes. We observe that the difference of average cost across the three traveler types is lower in het, het_sc, hom_sc, and zero, in comparison to curr. Moreover, we observe that for all type of travelers, the average travel cost is lower in het, het_sc, hom_sc, and zero, in comparison to curr[9]. Furthermore, we note that this observation not only holds in the averaged sense but also in a distributional sense[10]. Thus, curr may not be preferred by any type of traveler.

Next, in Figure 5.11, we compare different pricing schemes in terms of two metrics: average travel time and the distribution-welfare metric (as defined in (5.8)-(i)). Our results show that all pricing schemes, except for hom, outperform curr on both metrics. Additionally, we present a Pareto front (dotted line) that illustrates the trade-off between minimizing average travel time and reducing distributional disparities. The method used to compute this trade-off curve is detailed in [129].

Revenue considerations

Another important aspect of determining the congestion pricing scheme is the revenue it generates, which could be used for maintenance or expansion of existing transportation infrastructure, enhancing public transportation options, amongst other things. Figure 5.12

[9]The pricing scheme hom results in higher travel cost because it cannot differentiate between the types of traveler and charges higher tolls to travelers in order to ensure minimum average travel time.

[10]This result is illustrated in [129, Table 2-4], which presents the proportion of travelers of a particular type experiencing travel costs surpassing a predetermined threshold. We observe that, regardless of the value of threshold and the type of travelers, the proportion of travelers experiencing cost higher than a threshold is higher in curr in comparison to het, het_sc, hom_sc, and zero.

presents a comparison of different congestion pricing scheme in terms of total revenue. As per data released by the Metropolitan Transportation Commission (MTC) [11] a total toll revenue of $\$633,932,206$ was collected in the Bay Area during the year 2019-2020. Our calibrated model in curr predicts toll revenues on the same order of magnitude but slightly lower than MTC data. The mismatch between our prediction and MTC data is attributed to the fact that *(i)* our analysis only focuses on the morning rush hour, but MTC data also includes tolls collected beyond the morning rush hour as well, *(ii)* MTC data also includes tolls on HOV (High Occupancy Vehicle) lanes which are currently not added in our analysis, *(iii)* there is some additional demand incoming from other nearby cities not included in our analysis, and *(iv)* higher tolls are charged to multi-axle vehicles, with tolls charged as high as $36 in 2019.[12]

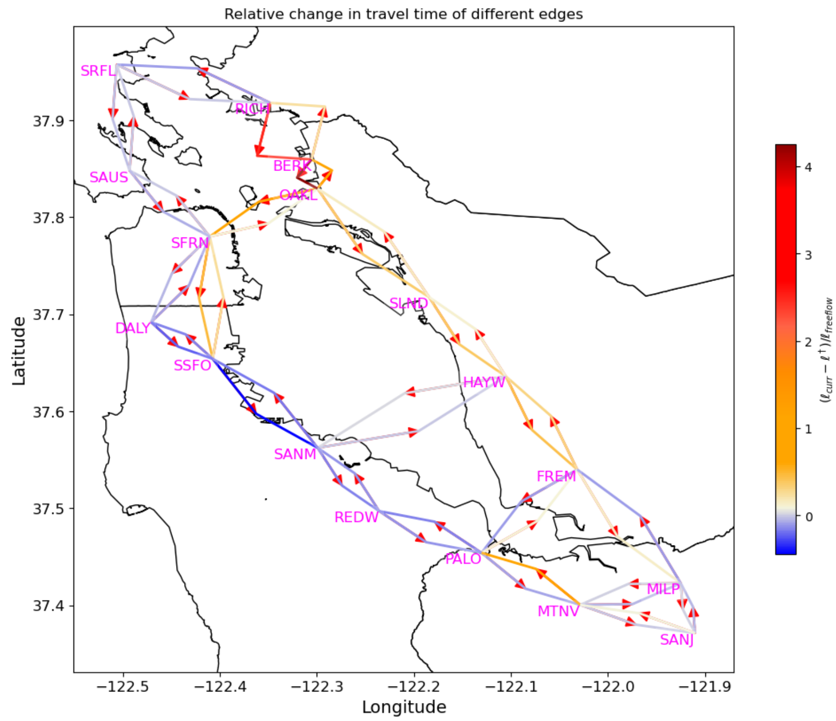
Notably, hom generates the highest revenue as it applies uniformly higher prices across all edges, irrespective of traveler types, with the goal of achieving a minimum average travel time. Moreover, the revenue of the other three pricing schemes hom_sc, het and het_sc are comparable to that of curr with het being slightly higher and hom_sc and het_sc being slightly lower.

5.7 Discussion

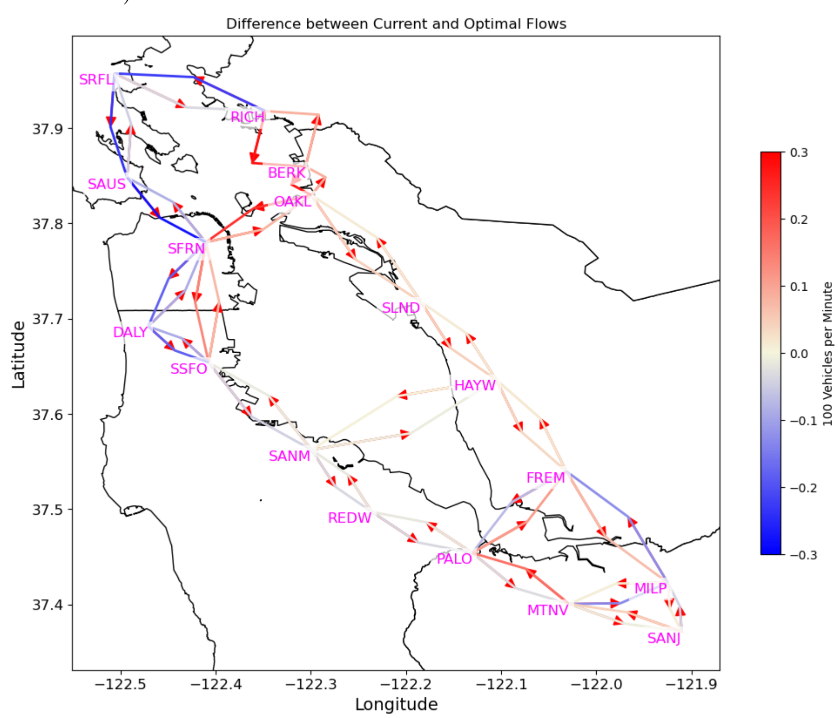
We study the problem of designing congestion pricing schemes that minimize the total travel time while reducing the distributional impacts of charging prices to travelers with heterogeneous value-of-time. We propose a linear programming-based approach to design four types of pricing schemes, varying based on (a) whether tolls are based on value-of-time and (b) whether all or only some network edges are tolled. Evaluating these schemes on the San Francisco Bay Area highway network shows that they outperform the current system in reducing congestion, improving the cost distribution, and generating revenue. Heterogeneous pricing schemes, in particular, offer travel cost distributions that are less disparate, highlighting their potential for future research. Key research directions include exploring rebate programs to support heterogeneous pricing, expanding the analysis to include incoming travelers from other cities, and considering mode choice and demand elasticity from factors like remote work.

[11] available at `https://mtc.ca.gov/about-mtc/authorities/bay-area-toll-authority/historic-toll-paid-vehicle-counts-toll-revenue`

[12] refer to `http://tinyurl.com/MTC-Multi-Axle`)

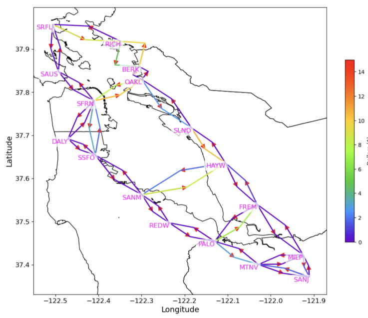


(a) Proportional travel time increase under curr (normalized by free flow travel time).

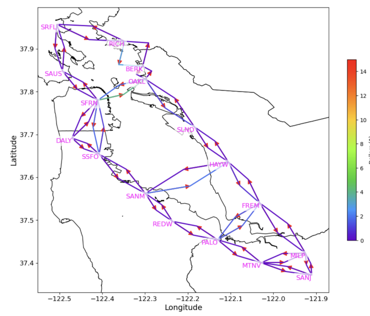


(b) Difference between equilibrium flow induced by curr and optimal flow.

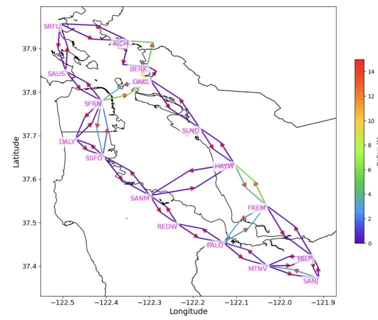
Figure 5.7: Current congestion pricing scheme



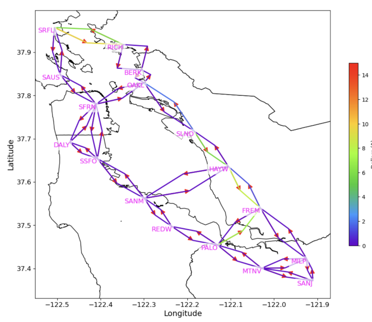
(a) Tolls under hom



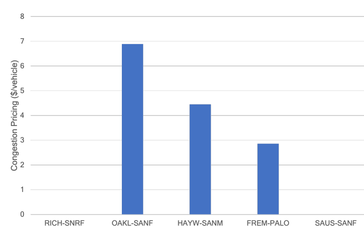
(b) Tolls under het for low type



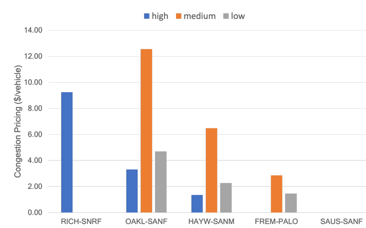
(c) Tolls under het for middle type



(d) Tolls under het for high type



(e) Tolls under hom_sc



(f) Tolls under het_sc

Figure 5.8: Toll values under congestion pricing schemes hom, het, hom_sc, and het_sc.

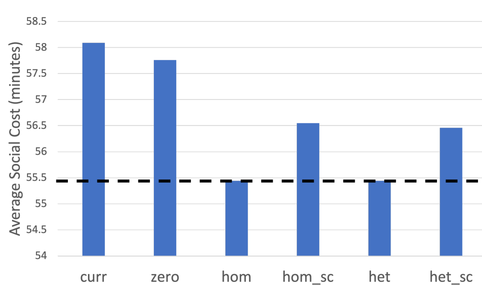


Figure 5.9: Comparison of average social cost per traveler for curr, zero, hom, hom_sc, het, and het_sc. Here, the dashed line represents the optimal value of average travel time computed by solving (5.5).

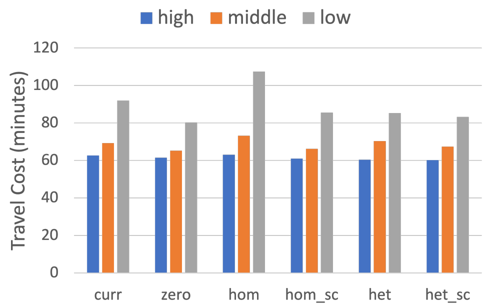


Figure 5.10: Average travel cost experienced by different types of travelers under different tolling schemes.

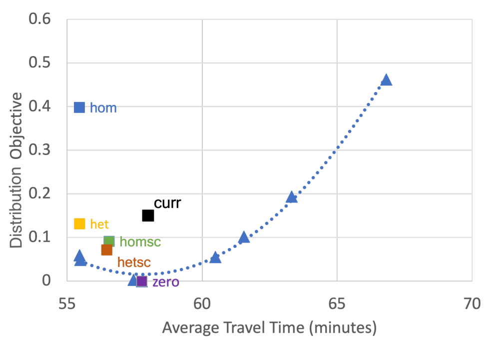


Figure 5.11: Trade-off between average travel time and cost distribution: The blue triangles represent different pricing schemes, positioned near the Pareto curve (a polynomial best-fit curve through the triangle points), based on computations detailed in [129].

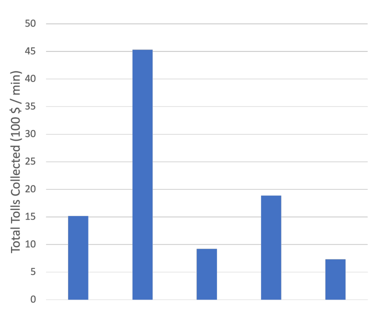


Figure 5.12: Comparison of total revenue collected for current, hom, hom-sc, het, het-sc.

Chapter 6

Social Choice with Changing Preferences

When Moses approached the camp and saw the calf and the dancing, his anger burned and he threw the tablets out of his hands, breaking them to pieces at the foot of the mountain.

Exodus 32:19-20

This chapter can be found in [107].

6.1 Introduction

Social choice theory [11] is a classic subfield of economics and philosophy that seeks to identify decisions that a social planner may make for a group based on the preferences of the group's members. In particular, the standard theorems of social choice are so-called *representation theorems* that provide constraints on the kinds of social alternatives that may be chosen given axioms on individual preferences. However, social choice generally operates in an environment where preferences are static and are not shaped by previous decisions that have been made by the social planner [75]. This leads to various critiques of social choice, for example from Pettigrew, who argues that utility functions of individuals *change over time* [162], and Parkes and Procaccia, who study the problem of group decision making with changing preferences from a voting-theoretic perspective [155]. In recent years, there has also been work on studying dynamic preferences in other contexts [46] [147].

In parallel, there is a standard theory of dynamic decision-making in the Markov Decision Process (MDP) literature [165] that studies memoryless state-transition models with reward functions and policies that maximize long-run rewards. MDPs are used in many decision-making tasks, most commonly in reinforcement learning applications [20].

As automated decision-making systems begin to make a larger fraction of choices for groups of individuals, we believe that a theory which joins the rigorous representation theorems of social choice theory with the dynamic nature of sequential decision-making is required. We aim to provide an optimality criterion for policies in such a setting by drawing on the rich existing work in social choice theory.

Our Results and Contributions

Our main contribution is to use a particular model of dynamic social choice, called a *Social Choice MDP* [155], and to show how we can draw on representation theorems to constrain reward functions and optimal policies to agree with the social welfare functional of Utilitarianism. In order to prove this representation theorem, we connect the class of reward functions of the Social Choice MDP with the individual utility functions of the agents in the group by providing axiomatic constraints which are necessary and sufficient for the reward function of the Social Choice MDP to agree with Utilitarianism. We also characterize the class of policies that arise as a result of implementing the long-run maximization of (Quasi-)Utilitarian rewards, and show that these policies lead to reasonable optimality criteria. This leads to axiomatic constraints on the value function of the optimization problem. Finally, we note that there are axioms that are standard in social choice theory, but whose validity breaks down in the dynamic setting. The most prominent of these is the (local) version of the Pareto axiom [155].

Previous Work

There has been previous work on studying dynamic social choice with evolving preferences by Parkes and Procaccia [155]. The main difference between this previous work and our contribution is that we use a different way to map social choice concepts to MDPs. While Parkes and Procaccia focus on axioms on social choice functions as constraints on policies, we focus on axioms on social welfare functionals as constraints on the reward function. The significance of this difference is that we can draw on the rich work in social choice theory on representation theorems for social welfare functionals, especially representation theorems for Utilitarianism [23]. Relatedly, while Parkes and Procaccia assume that group members have only ordinal preferences, we assume that their preferences are represented by (cardinal) utility functions.

6.2 Social Choice Theory

In this section, we introduce the basics of (classic) social choice theory [11, 75]. We start with a non-empty, finite set V of *group members* and a non-empty, finite set X of *social* *alternatives*. Let $\mathcal{U}(X)$ be $\{u \mid u : X \to \mathbb{R}\}$, the set of all utility functions over the social alternatives. Then, we define:

Definition 6.2.1. *A* profile *is a function* $U : V \to \mathcal{U}(X)$.

This means that a profile is an assignment of utility functions to group members. For every $i \in V$, we write $U_i(x)$ as shorthand for $U(i)(x)$. Let $\mathscr{U}$ be the set of all profiles.

In social choice theory, we are interested in how a group, or a 'social planner', should make decisions based on the preferences of all group members. There are different ways of formalizing this question. First, we can study social choice functions:

Definition 6.2.2. *A* social choice function (SCF) *is a map* $f : \mathcal{D} \to X$*, where* $\mathcal{D}$ *is some set of profiles.*

Given some profile $U \in dom(f)$, a SCF f selects a preferred social alternative $f(U) \in X$. Note that a SCF f must select a unique $x \in X$ for each $U \in \mathcal{D}$. This is a potential drawback of SCFs, as there may be situations in which different alternatives are equally good. In this case, SCFs require the introduction of arbitrary tie-breakers. Further, SCFs do not encode any information about the ranking among the social alternatives which are not chosen. We can avoid both of these problems by focusing instead on social welfare functions:

Definition 6.2.3. *A* social welfare functional (SWF) *is a map* $f : \mathcal{D} \to \mathcal{B}(X)$*, where* $\mathcal{D}$ *is some set of profiles and* $\mathcal{B}(X)$ *is the set of all binary relations on* X*.*

Given some profile $U \in dom(f)$, a SWF f returns a binary relation on X, which we interpret as a 'social preference relation'. For any profile U, we write $xf(U)y$ if $(x, y) \in f(U)$. The intended interpretation of $xf(U)y$ is that 'x is socially preferred to y'. We write $xP(f(U))y$ if $xf(U)y$ and not $yf(U)x$. The intended interpretation of $xP(f(U))y$ is that 'x is strictly socially preferred to y'.

Much work in social choice theory focuses on axioms which are imposed either on SCFs or SWFs (for example, some form of the Pareto principle). Here are Pareto axioms for SCFs and SWFs:

> SCF f satisfies **Pareto (SCF)** if for all $U \in dom(f)$ and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $f(U) \neq y$.
>
> SWF f satisfies **Pareto (SWF)** if for all $U \in dom(f)$ and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $xP(f(U))y$.

Research in social choice theory focuses in particular on *representation theorems*: finding a set of axioms which are necessary and sufficient for a SCF or SWF to be representable by a certain functional form [26]. An example is the SWF of *Utilitarianism*:

Definition 6.2.4. *SWF* f *is* Utilitarianism *if for all* $U \in dom(f)$*,* $x, y \in X$*,*

$$xf(U)y \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y).$$

In English, an alternative x is socially preferred to y if the sum of utilities of all the members of the group in x is bigger than the sum of utilities of the members in y. Consider the following axioms:

> SWF f satisfies **Universal Domain** if $dom(f)$ is the set of all profiles.

> SWF f satisfies **Transitivity** (**Completeness**) if for all $U \in dom(f)$, $f(U)$ is transitive (complete).

> SWF f satisfies **Independence of Irrelevant Alternatives (IIA)** if for all $U, U' \in dom(f)$ and $x, y \in X$, if $U_i(x) = U'_i(x)$ and $U_i(y) = U'_i(y)$ for all $i \in V$, then $xf(U)y$ if and only if $xf(U')y$.

Definition 6.2.5. *Two profiles U and U' satisfy* cardinal unit comparability, *written $U \sim_{CUC} U'$, if there is a $\beta \in \mathbb{R}$ with $\beta > 0$ and for every $i \in V$, there is some $\alpha_i \in \mathbb{R}$ such that for all $x \in X$, $U_i(x) = \alpha_i + \beta U'_i(x)$.*

> SWF f satisfies **CUC-Invariance** if for all $U, U' \in dom(f)$, if $U \sim_{CUC} U'$, then $f(U) = f(U')$.

> SWF f satisfies **Functional Anonymity** if for all $U, U' \in dom(f)$ and permutations $\rho : V \to V$, if for all $i \in V$, $U'_i = U_{\rho(i)}$, then $f(U) = f(U')$.

It has been shown that these axioms characterize Utilitarianism:

Theorem 6.2.6. *A SWF f is Utilitarianism if and only if f satisfies **Universal Domain**, **Transitivity**, **Completeness**, **IIA**, **Pareto (SWF)**, **CUC-Invariance** and **F-Anonymity** [48].*

6.3 Dynamic Decision-Making for Groups

Social choice theory normally considers static decision-making for groups. While this is amenable to analysis (through representation theorems), there is a critique of social choice in that it does not consider the case when preferences shift over time. Here, we consider the dynamic setting where the group members' preferences are changing over time according to a probabilistic model that is known to the social planner.

Markov Decision Processes

This section introduces our model for dynamic decision-making for groups. We consider Markov Decision Processes, which are memoryless state-transition models along with a reward function, which we define as:

Definition 6.3.1. *A* Markov Decision Process *is a tuple $\langle \mathcal{S}, \mathcal{A}, R, P \rangle$ where $\mathcal{S}$ is a finite non-empty set, $\mathcal{A}$ is a finite non-empty set, $P : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is a probability function and $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is a reward function.*

Further, the probability function satisfies the *Markov assumption*, which means that the probability of the next state only depends on the current state-action pair. We also define the notion of a policy:

Definition 6.3.2. *A (deterministic)* policy *is a function $\pi : \mathcal{S} \to \mathcal{A}$.*

The Social Choice MDP Model

Given the static models from social choice and the dynamic, state-transition based models from the MDP literature, we seek to define a model for dynamic decision-making when group members' preferences are shifting over time in response to actions taken by the social planner.

Definition 6.3.3. *A* Social Choice Markov Decision Process *is a tuple $\langle \mathcal{S}, \mathcal{A}, R, P \rangle$ where $\mathcal{S}$ is a non-empty finite set of profiles $U : V \to \mathcal{U}(X)$, $\mathcal{A}$ is the set of finite social alternatives X, $P : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is a probability function and $R : \mathscr{U} \times \mathcal{A} \to \mathbb{R}$ is a function.*

Our model differs from other Social Choice MDP models in two ways: first, we assume cardinal preferences, which give rise to MDP state spaces that are comprised of assignments of utility functions to group members, and second, our reward functions are defined over $\mathscr{U}$, the set of all profiles, in order for our representation theorems to hold.

From SCFs to SWFs

Note that in a Social Choice MDP, a policy $\pi : \mathcal{S} \to \mathcal{A}$ is a social choice function $f : \mathcal{D} \to X$, since $\mathcal{S}$ is a set of profiles and $\mathcal{A}$ is a set of social alternatives. Thus, [155] apply insights from the social choice literature to characterize policies in Social Choice MDPs. However, observe that there is also a correspondence between reward functions and SWFs. In particular, every reward function R in a Social Choice MDP induces a social welfare functional $f_R : \mathcal{D} \to \mathcal{B}(X)$.

Definition 6.3.4. *Given a reward function R, we define the corresponding SWF f_R for every profile U and all $x, y \in X$:*

$$x f_R(U) y \iff R(U, x) \geq R(U, y).$$

We can use this correspondence to use social choice axioms on SWFs as constraints on the reward function. One natural choice is the Utilitarian reward function: for every $U \in \mathcal{S}$ and $a \in \mathcal{A}$, $R(U, a) = \sum_{i \in V} U_i(a)$. However, instead of requiring the reward function to be strictly Utilitarian, we can also focus on the weaker requirement that it must agree with Utilitarianism up to strictly increasing transformations:

Definition 6.3.5. *A reward function* $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ *is* Quasi-Utilitarian *if for every* $U \in \mathcal{S}$ *and* $a \in \mathcal{A}$, $R(U, a) = f\left(\sum_{i \in V} U_i(a)\right)$, *where* $f : \mathbb{R} \to \mathbb{R}$ *is a strictly increasing function.*

6.4 Quasi-Utilitarian Characterization

In this section, we introduce constraints on the reward function R which entail that R agrees with the Utilitarianism social welfare functional and give a characterization of the policies generated by the Quasi-Utilitarian reward function.

Reward functions and SWFs

Using the mapping from reward functions and SWFs introduced above, we will impose the following axioms on the reward function R, which correspond to axioms on the induced SWF f_R:

> Reward function R satisfies **Pareto (SWF)** if for all $U \in dom(f_R)$ and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $xP(f_R(U))y$.
>
> Reward function R satisfies **Independence of Irrelevant Alternatives (IIA)** if for all $U, U' \in dom(f_R)$ and $x, y \in X$, if $U_i(x) = U'_i(x)$ and $U_i(y) = U'_i(y)$ for all $i \in V$, then $x f_R(U) y$ if and only if $x f_R(U') y$.
>
> Reward function R satisfies **CUC-Invariance** if for all $U, U' \in dom(f_R)$, if $U \sim_{CUC} U'$, then $f_R(U) = f_R(U')$.
>
> Reward function R satisfies **Functional Anonymity** if for all $U, U' \in dom(f_R)$ and permutations $\rho : V \to V$, if for all $i \in V$, $U'_i = U_{\rho(i)}$, then $f_R(U) = f_R(U')$.

Then, we show in Appendix D:

Theorem 6.4.1. *The following are equivalent for any Social Choice MDP:*

1. *R satisfies* ***Pareto (SWF)***, ***IIA***, ***CUC-Invariance*** *and* ***Functional Anonymity***.
2. *R agrees with Utilitarianism, so for any profile U and* $x, y \in X$, *we have*

$$R(U, x) \geq R(U, y) \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y).$$

Equivalently, R is Quasi-utilitarian.

Long Run Maximization

To get from reward function to optimal policies, we need to make additional assumptions. In this section, we draw on standard results from the MDP literature to argue for a particular kind of policy.

Definition 6.4.2. *A* value function *is a map* $V : \Pi \times S \to \mathbb{R}$*, where* Π *is the set of all policies and* S *is the set of all states.*

Intuitively, $V(\pi, s)$ is the value of executing policy π starting in state s. Given a value function, we define:

Definition 6.4.3. *The policy* π^* *is* optimal relative to V *if for all states* $s \in S$*,* $\pi^* \in \arg\max_{\pi \in \Pi} V(\pi, s)$*.*

We assume that the value function satisfies the *Bellman equation* [165] for any $\pi \in \Pi$ and $s \in S$:

$$V(\pi, s) = R(s, \pi(s)) + \gamma \sum_{s' \in S} p(s' \mid s, \pi(s)) V(\pi, s'),$$

where $0 < \gamma < 1$. This means that the value of executing policy π starting in state s is the sum of the immediate reward $R(s, \pi(s))$ and the expected future value of executing π in the next state, discounted by γ.[1] Now we can appeal to a standard result in the theory of MDPs [180, 182]:

Theorem 6.4.4. *Let* $V : \Pi \times S \to \mathbb{R}$ *be a value function. Then the following are equivalent for any MDP:*

1. *V satisfies the Bellman equation.*

2. *V is the expected sum of discounted future rewards. So, for any π and s,*

$$V(\pi, s) = \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t R(s_t, \pi(s_t))\right],$$

   where s_t is a random variable describing the state after t steps starting in state s with policy π and the expectation is taken relative to the transition model P.

Taken together with theorem 6.4.1, we can use this result to characterize what we call the class of *long-run quasi-utilitarian* policies:

[1] There are interesting questions about how to choose the discount rate which we do not discuss here in detail, see e.g. [71].

Definition 6.4.5. *Given a social choice MDP, a policy* π^* *is* long-run quasi-utilitarian *if for all* $s \in \mathcal{S}$*,*

$$\pi^* \in \arg\max_{\pi \in \Pi} \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t f\left(\sum_{i \in V} U_i^t(a)\right)\right],$$

where U^t *is a random variable describing the profile after* t *steps starting in state* s *with policy* π*, the expectation is taken relative to the transition model* P*, and* $f : \mathbb{R} \to \mathbb{R}$ *is strictly increasing.*

We propose this as a reasonable optimality criterion for group decision making under changing preferences. We can characterize this class as follows:

Theorem 6.4.6. *Given a social choice MDP, assume that* V *satisfies* ***Bellman equation*** *and* R *satisfies* ***Weak Pareto****,* ***IIA****,* ***CUC-Invariance*** *and* ***Functional Anonymity****. Then, the following are equivalent for any policy* π*:*

1. π *is optimal relative to* V*,*
2. π *is long-run quasi-utilitarian.*

6.5 Discussion and Future Work

We finish by discussing some consequences of our approach to group decision making with changing preferences. As noted above, there are two different ways of mapping social choice concepts to MDPs. First, we can think of policies as social choice functions (SCFs) and use axioms on SCFs to constrain policies. Second, we can exploit a correspondence between reward functions and social welfare functionals (SWFs), which is our distinctive contribution. We also noted earlier that there are two versions of the Pareto axiom for SCFs and SWFs respectively. The axioms for group decision making with changing preferences we defend here imply that our reward function satisfies the Pareto axiom for the SWF induced by the reward function:

> Reward function R satisfies **Pareto (SWF)** if for all $U \in dom(f_R)$ and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $xP(f_R(U))y$.

However, the policies which satisfy our optimality criterion will *not*, in general, satisfy the Pareto axiom for SCFs:

> SCF f satisfies **Pareto (SCF)** if for all $U \in dom(f)$ and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $f(U) \neq y$.

Applied to policies π, this axiom states that for all profiles U and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $\pi(U) \neq y$. This means that if every group member assigns higher utility to social alternative x than to social alternative y, y will not be chosen by our policy. However, this will not be true in general. Suppose, for example, that y leads, with high probability, to a future trajectory of high reward, while x leads, with high probability, to a future trajectory of low reward. Then, a long-run optimal policy will often choose y over x even though all group members assign higher utility to x.

This is interesting, because Parkes and Proccacia seem to suggest that the latter version of the Pareto axiom is a normatively sound constraint on group decision making:

> "In the case of Pareto optimality, if at any point the members all prefer one choice to another then the latter choice should not be made by the organization." [155]

In our view, while Pareto optimality in this sense might perhaps be a compelling axiom in some social choice contexts, such as sequential voting, it is *not* compelling in the context of long-run welfare optimization. This shows that once we focus on a dynamic setting with changing preferences, some of the traditional axioms of social choice theory lose their justification. Thus, it is important to study group decision making with changing preferences in its own right.

Chapter 7

Conclusion

We consider several problems that arise when computational systems interact with and make decisions over the actions of strategic agents. We formulate problems in incentive design (in traffic routing and atomic aggregative games), causal discovery, coalition formation (in electric charging), congestion pricing, and social choice for groups of agents. In doing so, we propose novel algorithms for solving these problems and present numerical experiments and empirical studies that demonstrate their efficacy. As future work, we propose expanding the scope of the theoretical assumptions and the empirical studies in the chapter of this thesis to apply to real-time user data collected from a broad variety of settings.

Bibliography

[1] Daron Acemoglu and Martin Kaae Jensen. "Aggregate comparative statics". In: *Games and Economic Behavior* 81 (2013), pp. 27–49.

[2] Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. *Networks, shocks, and systemic risk*. Tech. rep. National Bureau of Economic Research, 2015.

[3] Jeffrey L Adler and Mecit Cetin. "A direct redistribution model of congestion pricing". In: *Transportation research part B: methodological* 35.5 (2001), pp. 447–460.

[4] Selin Damla Ahipaşaoğlu, Uğur Arıkan, and Karthik Natarajan. "Distributionally Robust Markovian Traffic Equilibrium". In: *Transportation Science* 53.6 (2019), pp. 1546–1562.

[5] Polina Alexeenko and Eilyan Bitar. "Achieving reliable coordination of residential plug-in electric vehicle charging: A pilot study". In: *Transportation Research Part D: Transport and Environment* 118 (2023), p. 103658.

[6] Electrify America. *How electric vehicle (EV) charging works*. 2021. URL: `%5Curl%7Bhttps://www.electrifyamerica.com/how-ev-charging-works/%7D`.

[7] Nir Andelman, Michal Feldman, and Yishay Mansour. "Strong price of anarchy". In: *Games and Economic Behavior* 65.2 (2009), pp. 289–317.

[8] Enrico Angelelli, Idil Arsik, Valentina Morandi, Martin Savelsbergh, and Maria Grazia Speranza. "Proactive route guidance to avoid congestion". In: *Transportation Research Part B: Methodological* 94 (2016), pp. 1–21.

[9] Enrico Angelelli, Valentina Morandi, Martin Savelsbergh, and Maria Grazia Speranza. "System optimal routing of traffic flows with user constraints using linear programming". In: *European journal of operational research* 293.3 (2021), pp. 863–879.

[10] Richard Arnott and Kenneth Small. "The economics of traffic congestion". In: *American scientist* 82.5 (1994), pp. 446–455.

[11] Kenneth Arrow. *Social Choice and Individual Values*. John Wiley & Sons, 1951.

[12] Peiman Asadi, Sebastian Engelke, and Anthony C. Davison. "Optimal Regionalization of Extreme Value Distributions for Flood Estimation". In: *Journal of Hydrology* 556 (2018), pp. 182–193.

[13] I.C. Athira, C.P. Muneera, K. Krishnamurthy, and M.V.L.R. Anjaneyulu. "Estimation of Value of Travel Time for Work Trips". In: *Transportation Research Procedia* 17 (2016). International Conference on Transportation Planning and Implementation Methodologies for Developing Countries (12th TPMDC) Selected Proceedings, IIT Bombay, Mumbai, India, 10-12 December 2014, pp. 116–123. ISSN: 2352-1465. DOI: `https://doi.org/10.1016/j.trpro.2016.11.067`. URL: `https://www.sciencedirect.com/science/article/pii/S2352146516306810`.

[14] Robert J Aumann. "Acceptable points in general cooperative n-person games". In: *Contributions to the Theory of Games* 4.40 (1959), pp. 287–324.

[15] Yoram Bachrach, Vasilis Syrgkanis, Éva Tardos, and Milan Vojnović. "Strong price of anarchy, utility games and coalitional dynamics". In: *Algorithmic Game Theory: 7th International Symposium, SAGT 2014, Haifa, Israel, September 30–October 2, 2014. Proceedings 7*. Springer. 2014, pp. 218–230.

[16] Yu Bai, Chi Jin, Huan Wang, and Caiming Xiong. "Sample-efficient learning of stackelberg equilibria in general-sum games". In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 25799–25811.

[17] Jean-Bernard Baillon and Roberto Cominetti. "Markovian Traffic Equilibrium". In: *Mathematical Programming* (Feb. 2008).

[18] Ian C Barnes, Karen Trapenberg Frick, Elizabeth Deakin, and Alexander Skabardonis. "Impact of peak and off-peak tolls on traffic in san francisco–oakland bay bridge corridor in california". In: *Transportation research record* 2297.1 (2012), pp. 73–79.

[19] Jorge Barrera and Alfredo Garcia. "Dynamic incentives for congestion control". In: *IEEE Transactions on Automatic Control* 60.2 (2014), pp. 299–310.

[20] Andrew G. Barto, R. S. Sutton, and C. J. C. H. Watkins. "Learning and Sequential Decision Making". In: *Learning and computational Neuroscience*. MIT Press, 1989, pp. 539–602.

[21] Tamer Başar. "Affine incentive schemes for stochastic systems with dynamic information". In: *SIAM Journal on Control and Optimization* 22.2 (1984), pp. 199–210.

[22] Martin Beckmann, Charles B McGuire, and Christopher B Winsten. *Studies in the Economics of Transportation*. Tech. rep. 1956.

[23] Jeremy Bentham. *An Introduction to the Principles of Morals and Legislation*. 1789.

[24] David Bernstein. "Congestion pricing with tolls and subsidies". In: *Pacific Rim TransTech Conference—Volume II: International Ties, Management Systems, Propulsion Technology, Strategic Highway Research Program*. ASCE. 1993, pp. 145–151.

[25] Dimitris Bertsimas, Vishal Gupta, and Ioannis Ch Paschalidis. "Data-driven estimation in equilibrium using inverse optimization". In: *Mathematical Programming* 153 (2015), pp. 595–633.

[26] Charles Blackorby, Walter Bossert, and David Donaldson. "Utilitarianism and the theory of justice". In: *Handbook of social choice and welfare* 1 (2002), pp. 543–596.

[27] NEF Bloomberg. "Hitting the EV inflection Point". In: *Electric Vehicle Price Parity Phasing Out Combustion Vehicle Sales in Europe; Bloomberg: New York, NY, USA* (2021).

[28] Avrim Blum, Nika Haghtalab, and Ariel D Procaccia. "Learning optimal commitment to overcome insecurity". In: *Advances in Neural Information Processing Systems* 27 (2014).

[29] Vincenzo Bonifaci, Mahyar Salek, and Guido Schäfer. "Efficiency of restricted tolls in non-atomic network routing games". In: *Algorithmic Game Theory: 4th International Symposium, SAGT 2011, Amalfi, Italy, October 17-19, 2011. Proceedings 4*. Springer. 2011, pp. 302–313.

[30] Vivek S Borkar. "Stochastic approximation with two time scales". In: *Systems & Control Letters* 29.5 (1997), pp. 291–294.

[31] Vivek S Borkar. *Stochastic approximation: a dynamical systems viewpoint*. Vol. 48. Springer, 2009.

[32] Vivek S Borkar and Sarath Pattathil. "Concentration bounds for two time scale stochastic approximation". In: *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2018, pp. 504–511.

[33] Yann Bramoullé, Andrea Galeotti, and Brian W Rogers. *The Oxford handbook of the economics of networks*. Oxford University Press, 2016.

[34] Yann Bramoullé and Rachel Kranton. "Public goods in networks". In: *Journal of Economic theory* 135.1 (2007), pp. 478–494.

[35] Luce Brotcorne, Martine Labbé, Patrice Marcotte, and Gilles Savard. "A bilevel model for toll optimization on a multicommodity transportation network". In: *Transportation science* 35.4 (2001), pp. 345–358.

[36] Abby Brown, Jeff Cappellucci, Alexia Heinrich, and Emma Cost. *Electric Vehicle Charging Infrastructure Trends from the Alternative Fueling Station Locator: Third Quarter 2023*. Tech. rep. National Renewable Energy Laboratory (NREL), Golden, CO (United States), 2024.

[37] Philip N Brown and Jason R Marden. "A study on price-discrimination for robust social coordination". In: *2016 American Control Conference (ACC)*. IEEE. 2016, pp. 1699–1704.

[38] Yilan Chen, Daniel E Ochoa, Jason R Marden, and Jorge I Poveda. "High-Order Decentralized Pricing Dynamics for Congestion Games: Harnessing Coordination to Achieve Acceleration". In: *2023 American Control Conference (ACC)*. IEEE. 2023, pp. 1086–1091.

[39] Steve Chien and Alistair Sinclair. “Strong and pareto price of anarchy in congestion games”. In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2009, pp. 279–291.

[40] Chih-Yuan Chiu, Kshitij Kulkarni, and Shankar Sastry. “Towards Dynamic Causal Discovery with Rare Events: A Nonparametric Conditional Independence Test”. In: *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE. 2023, pp. 7610–7616.

[41] Riccardo Colini-Baldeschi, Roberto Cominetti, Panayotis Mertikopoulos, and Marco Scarsini. “When is selfish routing bad? The price of anarchy in light and heavy traffic”. In: *Operations Research* 68.2 (2020), pp. 411–434.

[42] Roberto Cominetti, Valerio Dose, and Marco Scarsini. “The price of anarchy in routing games as a function of the demand”. In: *Mathematical Programming* (2021), pp. 1–28.

[43] Giacomo Como and Rosario Maggistro. “Distributed Dynamic Pricing of Multiscale Transportation Networks”. In: *IEEE Transactions on Automatic Control* (2021).

[44] Giacomo Como and Rosario Maggistro. “Distributed Dynamic Pricing of Multiscale Transportation Networks”. In: *IEEE Transactions on Automatic Control* 67.4 (2022), pp. 1625–1638.

[45] Lauren Craik and Hamsa Balakrishnan. “Equity impacts of the London congestion charging scheme: an empirical evaluation using synthetic control methods”. In: *Transportation research record* 2677.5 (2023), pp. 1017–1029.

[46] Elisabeth Crawford and Manuela Veloso. “Learning dynamic preferences in multi-agent meeting scheduling”. In: *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*. 2005, pp. 487–490.

[47] Qiwen Cui, Maryam Fazel, and Simon S Du. “Learning Optimal Tax Design in Nonatomic Congestion Games”. In: *arXiv preprint arXiv:2402.07437* (2024).

[48] Claude D’Aspremont and Louis Gevers. “Equity and the Informational Basis of Collective Choice”. In: *The Review of Economic Studies* 44.2 (1977), pp. 199–209. DOI: 10.2307/2297061.

[49] Stella Dafermos. “Sensitivity analysis in variational inequalities”. In: *Mathematics of Operations Research* 13.3 (1988), pp. 421–434.

[50] Carlos F Daganzo. “A pareto optimum congestion reduction scheme”. In: *Transportation Research Part B: Methodological* 29.2 (1995), pp. 139–154.

[51] Luca Deori, Kostas Margellos, and Maria Prandini. “On the connection between Nash equilibria and social optima in electric vehicle charging control games”. In: *IFAC-PapersOnLine* 50.1 (2017), pp. 14320–14325.

[52] Lydia Depillis, Rebecca Lieberman, and Crista Chapman. *How the costs of car ownership add up*. Oct. 2023. URL: https://www.nytimes.com/interactive/2023/10/07/business/car-ownership-costs.html.

[53] Nico Dogterom, Dick Ettema, and Martin Dijst. "Tradable credits for managing car travel: a review of empirical research and relevant behavioural approaches". In: *Transport Reviews* 37.3 (2017), pp. 322–343.

[54] US DOT. *Income-Based Equity Impacts of Congestion Pricing*. 2008.

[55] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. "Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator". In: *The Annals of Mathematical Statistics* 27.3 (1956), pp. 642–669. DOI: `10.1214/aoms/1177728174`.

[56] Joakim Ekström, Leonid Engelson, and Clas Rydergren. "Heuristic algorithms for a second-best congestion pricing problem". In: *NETNOMICS: Economic Research and Electronic Networking* 10 (2009), pp. 85–102.

[57] Jonas Eliasson. "Road pricing with limited information and heterogeneous users: A successful case". In: *The annals of regional science* 35 (2001), pp. 595–604.

[58] Jonas Eliasson and Lars-Göran Mattsson. "Equity effects of congestion pricing: quantitative methodology and a case study for Stockholm". In: *Transportation Research Part A: Policy and Practice* 40.7 (2006), pp. 602–620.

[59] Sebastian Engelke and Stanislav Volgushev. "Structure Learning for Extremal Tree Models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* (2022).

[60] Doris Entner and Patrik O Hoyer. "On Causal Discovery from Times Series Data Using FCI". In: *Probabilistic graphical models* (2010), pp. 121–128.

[61] Amir Epstein, Michal Feldman, and Yishay Mansour. "Strong equilibrium in cost sharing connection games". In: *Proceedings of the 8th ACM conference on Electronic commerce*. 2007, pp. 84–92.

[62] J Joaquin Escudero-Garzas and Gonzalo Seco-Granados. "Charging station selection optimization for plug-in electric vehicles: An oligopolistic game-theoretic framework". In: *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*. IEEE. 2012, pp. 1–8.

[63] EVgo. *What is EVgo — Electric Vehicle (EV) Fast Charging Stations*. `https://www.evgo.com/ev-drivers/the-evgo-network/`. 2021.

[64] Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer, 2003.

[65] Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[66] Michal Feldman and Ophir Friedler. "A unified framework for strong price of anarchy in clustering games". In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2015, pp. 601–613.

[67] Rui Feng, Huixia Zhang, Bin Shi, Qian Zhong, and Baozhen Yao. "Collaborative road pricing strategy for heterogeneous vehicles considering emission constraints". In: *Journal of Cleaner Production* 429 (2023), p. 139561.

[68] Bryce L Ferguson, Dario Paccagnan, Bary SR Pradelski, and Jason R Marden. "Collaborative Coalitions in Multi-Agent Systems: Quantifying the Strong Price of Anarchy for Resource Allocation Games". In: *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE. 2023, pp. 3238–3243.

[69] Paolo Ferrari. "Road network toll pricing and social welfare". In: *Transportation Research Part B: Methodological* 36.5 (2002), pp. 471–483.

[70] Lisa Fleischer, Kamal Jain, and Mohammad Mahdian. "Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games". In: *45th Annual IEEE Symposium on Foundations of Computer Science*. IEEE. 2004, pp. 277–285.

[71] Marc Fleurbaey and Stephane Zuber. "Climate Policies Deserve a Negative Discount Rate". In: *Chicago Journal of International Law* 13.2 (2013).

[72] Karen T Frick, Steve Heminger, and Hank Dittmar. "Bay bridge congestion-pricing project: Lessons learned to date". In: *Transportation Research Record* 1558.1 (1996), pp. 29–38.

[73] Zhengtang Fu, Peiwu Dong, and Yanbing Ju. "An intelligent electric vehicle charging system for new energy companies based on consortium blockchain". In: *Journal of Cleaner Production* 261 (2020), p. 121219.

[74] Drew Fudenberg and David Levine. *The theory of learning in games*. Vol. 2. MIT press, 1998.

[75] Wulf Gaertner. *A Primer in Social Choice Theory*. Oxford University Press, 2006.

[76] Lingwen Gan, Ufuk Topcu, and Steven H Low. "Optimal decentralized protocol for electric vehicle charging". In: *IEEE Transactions on Power Systems* 28.2 (2012), pp. 940–951.

[77] Nicola Gatti, Marco Rocco, and Tuomas Sandholm. "On the verification and computation of strong Nash equilibrium". In: *arXiv preprint arXiv:1711.06318* (2017).

[78] Kurtuluş Gemici, Elias Koutsoupias, Barnabé Monnot, Christos Papadimitriou, and Georgios Piliouras. "Wealth inequality and the price of anarchy". In: *arXiv preprint arXiv:1802.09269* (2018).

[79] Clark Glymour, Kun Zhang, and Peter Spirtes. "Review of Causal Discovery Methods Based on Graphical Models". In: *Frontiers in Genetics* 10 (2019).

[80] Nicola Gnecco, Nicolai Meinshausen, Jonas Peters, and Sebastian Engelke. "Causal Discovery in Heavy-tailed Models". In: *The Annals of Statistics* 49.3 (2021), pp. 1755–1778.

[81] Eric J Gonzales and Eleni Christofa. “Empirical assessment of bottleneck congestion with a constant and peak toll: San Francisco–Oakland Bay Bridge”. In: *EURO Journal on Transportation and Logistics* 3.3 (2015), pp. 267–288.

[82] PB Goodwin. “How to make road pricing popular”. In: *Economic Affairs* 10.5 (1990), pp. 6–7.

[83] Phil B Goodwin. “The rule of three: a possible solution to the political problem of competing objectives for road pricing”. In: *Traffic engineering & control* 30.10 (1989), pp. 495–497.

[84] Sergio Grammatico. “Dynamic control of agents playing aggregative games with coupling constraints”. In: *IEEE Transactions on Automatic Control* 62.9 (2017), pp. 4537–4548.

[85] Clive William John Granger. “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods”. In: *Econometrica*. Vol. 37. 1969, pp. 424–438.

[86] Giambattista Gruosso. “Analysis of impact of electrical vehicle charging on low voltage power grid”. In: *2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)*. IEEE. 2016, pp. 1–6.

[87] Hugh Gunn. “Spatial and temporal transferability of relationships between travel demand, trip cost and travel time”. In: *Transportation Research Part E: Logistics and Transportation Review* 37.2-3 (2001), pp. 163–189.

[88] Xiaolei Guo and Hai Yang. “Pareto-improving congestion pricing and revenue refunding with multiple user classes”. In: *Transportation Research Part B: Methodological* 44.8-9 (2010), pp. 972–982.

[89] Libin Han, Chong Peng, and Zhenyu Xu. “The effect of commuting time on quality of life: evidence from China”. In: *International journal of environmental research and public health* 20.1 (2022), p. 573.

[90] Tobias Harks, Ingo Kleinert, Max Klimm, and Rolf H Möhring. “Computing network tolls with support constraints”. In: *Networks* 65.3 (2015), pp. 262–285.

[91] Tobias Harks and Julian Schwarz. “A unified framework for pricing in nonconvex resource allocation games”. In: *SIAM Journal on Optimization* 33.2 (2023), pp. 1223–1249.

[92] Christopher Harris. “On the rate of convergence of continuous-time fictitious play”. In: *Games and Economic Behavior* 22.2 (1998), pp. 238–259.

[93] Morris W Hirsch. “Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere”. In: *SIAM Journal on Mathematical Analysis* 16.3 (1985), pp. 423–439.

[94] Yu-Chi Ho, Peter B Luh, and Geert Jan Olsder. “A control-theoretic view on incentives”. In: *Automatica* 18.2 (1982), pp. 167–179.

[95] Martin Hoefer, Lars Olbrich, and Alexander Skopalik. "Taxing subnetworks". In: *International workshop on internet and network economics*. Springer. 2008, pp. 286–294.

[96] Josef Hofbauer and William H Sandholm. "On the global convergence of stochastic fictitious play". In: *Econometrica* 70.6 (2002), pp. 2265–2294.

[97] Olaf Jahn, Rolf H Möhring, Andreas S Schulz, and Nicolás E Stier-Moses. "System-optimal routing of traffic flows with user constraints in networks with congestion". In: *Operations research* 53.4 (2005), pp. 600–616.

[98] Devansh Jalota, Karthik Gopalakrishnan, Navid Azizan, Ramesh Johari, and Marco Pavone. "Online Learning for Traffic Routing under Unknown Preferences". In: *arXiv preprint arXiv:2203.17150* (2022).

[99] Devansh Jalota, Kiril Solovey, Karthik Gopalakrishnan, Stephen Zoepf, Hamsa Balakrishnan, and Marco Pavone. "When efficiency meets equity in congestion pricing and revenue refunding schemes". In: *Equity and Access in Algorithms, Mechanisms, and Optimization*. 2021, pp. 1–11.

[100] Devansh Jalota, Kiril Solovey, Matthew Tsao, Stephen Zoepf, and Marco Pavone. "Balancing fairness and efficiency in traffic routing via interpolated traffic assignment". In: *Autonomous Agents and Multi-Agent Systems* 37.2 (2023), p. 32.

[101] Kaushik Jana, Prajamitra Bhuyan, and Emma McCoy. "Causal Analysis at Extreme Quantiles with Application to London Traffic Flow Data". In: *ArXiv* (Aug. 2021).

[102] Vyacheslav V Kalashnikov, Roberto Carlos Herrera Maldonado, José-Fernando Camacho-Vallejo, and Nataliya I Kalashnykova. "A heuristic algorithm solving bilevel toll optimization problems". In: *The International Journal of Logistics Management* 27.1 (2016), pp. 31–51.

[103] George Karakostas and Stavros G Kolliopoulos. "Edge pricing of multicommodity networks for heterogeneous selfish users". In: *FOCS*. Vol. 4. 2004, pp. 268–276.

[104] Wajahat Khan, Furkan Ahmad, and Mohammad Saad Alam. "Fast EV charging station integration with grid ensuring optimal and quality power exchange". In: *Engineering Science and Technology, an International Journal* 22.1 (2019), pp. 143–152.

[105] Walid Krichene, Benjamin Drighès, and Alexandre Bayen. "On the Convergence of No-Regret Learning in Selfish Routing". In: *Proceedings of the 31st International Conference on Machine Learning*. Ed. by Eric P. Xing and Tony Jebara. Vol. 32. Proceedings of Machine Learning Research. Bejing, China: PMLR, June 2014, pp. 163–171.

[106] Sukanya Kudva, Kshitij Kulkarni, Chinmay Maheshwari, Anil Aswani, and Shankar Sastry. "Understanding the Impact of Coalitions between EV Charging Stations". In: *arXiv preprint arXiv:2404.03919* (2024).

[107] Kshitij Kulkarni and Sven Neth. "Social choice with changing preferences: Representation theorems and long-run policies". In: *arXiv preprint arXiv:2011.02544* (2020).

[108] Martine Labbé, Patrice Marcotte, and Gilles Savard. "A bilevel model of taxation and its application to optimal highway pricing". In: *Management science* 44.12-part-1 (1998), pp. 1608–1622.

[109] Rida Laraki and Panayotis Mertikopoulos. "Higher order game dynamics". In: *Journal of Economic Theory* 148.6 (2013), pp. 2666–2695.

[110] Torbjörn Larsson and Michael Patriksson. "Side constrained traffic equilibrium models—traffic management through link tolls". In: *Equilibrium and advanced transportation modelling*. Springer, 1998, pp. 125–151.

[111] S Lawphongpanich and Y Yin. "Pareto-improving congestion pricing for general road networks". In: *Technique Report. Gainesville, FL: Department of industrial and System Engineering, University of Florida* (2007).

[112] Siriphong Lawphongpanich and Donald W Hearn. "An MPEC approach to second-best toll pricing". In: *Mathematical programming* 101.1 (2004), pp. 33–55.

[113] Siriphong Lawphongpanich and Yafeng Yin. "Solving the Pareto-improving toll problem via manifold suboptimization". In: *Transportation Research Part C: Emerging Technologies* 18.2 (2010), pp. 234–246.

[114] Daniel A Lazar and Ramtin Pedarsani. "Optimal tolling for multitype mixed autonomous traffic networks". In: *IEEE Control Systems Letters* 5.5 (2020), pp. 1849–1854.

[115] Daniel A Lazar and Ramtin Pedarsani. "The Role of Differentiation in Tolling of Traffic Networks with Mixed Autonomy". In: *arXiv preprint arXiv:2103.13553* (2021).

[116] Phil LeBeau. "Traffic jams cost US $87 billion in lost productivity in 2018, and Boston and DC have the nation's worst". In: (2019).

[117] Woongsup Lee, Lin Xiang, Robert Schober, and Vincent WS Wong. "Electric vehicle charging stations with renewable power generators: A game theoretical analysis". In: *IEEE transactions on smart grid* 6.2 (2014), pp. 608–617.

[118] Erich Leo Lehmann. "Consistency and Unbiasedness of Certain Nonparametric Tests". In: *Annals of Mathematical Statistics* 22 (1951), pp. 165–179.

[119] Joshua Letchford, Vincent Conitzer, and Kamesh Munagala. "Learning and approximating the optimal strategy to commit to". In: *Algorithmic Game Theory: Second International Symposium, SAGT 2009, Paphos, Cyprus, October 18-20, 2009. Proceedings 2*. Springer. 2009, pp. 250–262.

[120] Dan Levin. "Taxation within Cournot oligopoly". In: *Journal of Public Economics* 27.3 (1985), pp. 281–290.

[121] Jianhui Li, Youcheng Niu, Shuang Li, Yuzhe Li, Jinming Xu, and Junfeng Wu. "Inducing Desired Equilibrium in Taxi Repositioning Problem with Adaptive Incentive Design". In: *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE. 2023, pp. 8075–8080.

[122] Jiayi Li, Matthew Motoki, and Baosen Zhang. "Socially Optimal Energy Usage via Adaptive Pricing". In: *arXiv preprint arXiv:2310.13254* (2023).

[123] Alvin Chong Lim. *Transportation network design problems: An MPEC approach*. The Johns Hopkins University, 2002.

[124] Xi Lin, Yafeng Yin, and Fang He. "Credit-based mobility management considering travelers' budgeting behaviors under uncertainty". In: *Transportation Science* 55.2 (2021), pp. 297–314.

[125] Nick Littlestone and Manfred K Warmuth. "The weighted majority algorithm". In: *Information and computation* 108.2 (1994), pp. 212–261.

[126] Boyi Liu, Jiayang Li, Zhuoran Yang, Hoi-To Wai, Mingyi Hong, Yu Marco Nie, and Zhaoran Wang. "Inducing Equilibria via Incentives: Simultaneous Design-and-Play Finds Global Optima". In: *arXiv preprint arXiv:2110.01212* (2021).

[127] Jennifer Liu. "Commuters in this city spend 119 hours a year stuck in traffic". In: (2020). Accessed: 2024-12-18. URL: `https://www.cnbc.com/2019/09/04/commuters-in-this-city-spend-119-hours-a-year-stuck-in-traffic.html`.

[128] Zhongjing Ma, Duncan S Callaway, and Ian A Hiskens. "Decentralized charging control of large populations of plug-in electric vehicles". In: *IEEE Transactions on control systems technology* 21.1 (2011), pp. 67–78.

[129] Chinmay Maheshwari, Kshitij Kulkarni, Druv Pai, Jiarui Yang, Manxi Wu, and Shankar Sastry. "Congestion Pricing for Efficiency and Equity: Theory and Applications to the San Francisco Bay Area". In: *arXiv preprint arXiv:2401.16844* (2024).

[130] Chinmay Maheshwari, Kshitij Kulkarni, Manxi Wu, and S. Shankar Sastry. "Dynamic Tolling for Inducing Socially Optimal Traffic Loads". In: *2022 American Control Conference (ACC)*. 2022, pp. 4601–4607.

[131] Chinmay Maheshwari, Kshitij Kulkarni, Manxi Wu, and S. Shankar Sastry. "Inducing Social Optimality in Games via Adaptive Incentive Design". In: *ArXiv 2204.05507* (2022).

[132] Chinmay Maheshwari, Kshitij Kulkarni, Manxi Wu, and Shankar Sastry. "Adaptive Incentive Design with Learning Agents". In: *arXiv preprint arXiv:2405.16716* (2024).

[133] Chinmay Maheshwari, Kshitij Kulkarni, Manxi Wu, and Shankar Sastry. "Dynamic Tolling for Inducing Socially Optimal Traffic Loads". In: *arXiv preprint arXiv:2110.08879* (2021).

[134] Chinmay Maheshwari, S Shankar Sastry, Lillian Ratliff, and Eric Mazumdar. "Follower Agnostic Methods for Stackelberg Games". In: *arXiv preprint arXiv:2302.01421* (2023).

[135] Daniel Malinsky and Peter L. Spirtes. "Causal Structure Learning from Multivariate Time Series in Settings with Unmeasured Confounding". In: *CD@KDD*. 2018.

[136] Traffic Assignment Manual. *For Application With a Large, High Speed Computer.* 1964.

[137] Patrice Marcotte and DL Zhu. "Existence and computation of optimal tolls in multiclass network equilibrium problems". In: *Operations Research Letters* 37.3 (2009), pp. 211–214.

[138] Negar Mehr and Roberto Horowitz. "Pricing traffic networks with mixed vehicle autonomy". In: *2019 American Control Conference (ACC)*. IEEE. 2019, pp. 2676–2682.

[139] Panayotis Mertikopoulos and Zhengyuan Zhou. "Learning in games with continuous action sets and unknown payoff functions". In: *Mathematical Programming* 173.1 (2019), pp. 465–507.

[140] David Meunier and Emile Quinet. "Value of time estimations in cost benefit analysis: the French experience". In: *Transportation Research Procedia* 8 (2015), pp. 62–71.

[141] Eduardo Mojica-Nava, Jorge I Poveda, and Nicanor Quijano. "Stackelberg Population Learning Dynamics". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE. 2022, pp. 6395–6400.

[142] Dov Monderer and Lloyd S Shapley. "Fictitious play property for games with identical interests". In: *Journal of economic theory* 68.1 (1996), pp. 258–265.

[143] Barnabé Monnot, Francisco Benita, and Georgios Piliouras. "How bad is selfish routing in practice?" In: *arXiv preprint arXiv:1703.01599* (2017).

[144] Michael Naaman. "On the Tight Constant in the Multivariate Dvoretzky–Kiefer–Wolfowitz Inequality". In: *Statistics & Probability Letters* 173 (2021), p. 109088.

[145] John H Nachbar. ""Evolutionary" selection dynamics in games: Convergence and limit properties". In: *International journal of game theory* 19.1 (1990), pp. 59–89.

[146] Katsuhiko Nakamura and Kara Maria Kockelman. "Congestion pricing and roadspace rationing: an application to the San Francisco Bay Bridge corridor". In: *Transportation Research Part A: Policy and Practice* 36.5 (2002), pp. 403–417.

[147] Sriraam Natarajan and Prasad Tadepalli. "Dynamic Preferences in Multi-Criteria Reinforcement Learning". In: ICML '05. Bonn, Germany, 2005, pp. 601–608. DOI: 10.1145/1102351.1102427.

[148] Rabia Nessah and Guoqiang Tian. "On the existence of strong Nash equilibria". In: *Journal of Mathematical Analysis and Applications* 414.2 (2014), pp. 871–885.

[149] Yu Marco Nie. “Transaction costs and tradable mobility credits”. In: *Transportation Research Part B: Methodological* 46.1 (2012), pp. 189–203.

[150] Steven J O’Hare, Richard D Connors, and David P Watling. “Mechanisms that govern how the price of anarchy varies with travel demand”. In: *Transportation Research Part B: Methodological* 84 (2016), pp. 55–80.

[151] Hassan Obeid, Ayse Tugba Ozturk, Wente Zeng, and Scott J Moura. “Learning and optimizing charging behavior at PEV charging stations: Randomized pricing experiments, and joint power and price optimization”. In: *Applied Energy* 351 (2023), p. 121862.

[152] Daniel E Ochoa and Jorge I Poveda. “High-performance optimal incentive-seeking in transactive control for traffic congestion”. In: *European Journal of Control* 68 (2022), p. 100696.

[153] Dario Paccagnan, Francesca Parise, and John Lygeros. “On the efficiency of Nash equilibria in aggregative charging games”. In: *IEEE control systems letters* 2.4 (2018), pp. 629–634.

[154] Raymond Palmquist, Daniel Phaneuf, and V Kerry Smith. *Measuring the values for time.* 2007.

[155] David C. Parkes and Ariel D. Procaccia. “Dynamic Social Choice with Evolving Preferences”. In: *In Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence* (2014).

[156] Michael Patriksson and R Tyrrell Rockafellar. “A mathematical model and descent algorithm for bilevel traffic management”. In: *Transportation Science* 36.3 (2002), pp. 271–291.

[157] Judea Pearl. *Causality: Models, Reasoning and Inference.* Cambridge University Press, 2009.

[158] Binghui Peng, Weiran Shen, Pingzhong Tang, and Song Zuo. “Learning optimal strategies to commit to”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 2019, pp. 2149–2156.

[159] Marco Percoco. “Heterogeneity in the reaction of traffic flows to road pricing: a synthetic control approach applied to Milan”. In: *Transportation* 42 (2015), pp. 1063–1079.

[160] Cora Beatriz Pérez-Ariza, Ann E. Nicholson, Kevin B. Korb, Steven Mascaro, and Chao Heng Hu. “Causal Discovery of Dynamic Bayesian Networks”. In: *Proceedings of the 25th Australasian Joint Conference on Advances in Artificial Intelligence*. AI’12. Sydney, Australia: Springer-Verlag, 2012, pp. 902–913. ISBN: 9783642351006.

[161] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms.* Adaptive Computation and Machine Learning. MIT Press, 2017. ISBN: 978-0-262-03731-0.

[162] Richard Pettigrew. *Choosing for Changing Selves*. Oxford University Press, 2020.

[163] Sock-Yong Phang and Rex S Toh. "Road congestion pricing in Singapore: 1975 to 2003". In: *Transportation Journal* (2004), pp. 16–25.

[164] Jorge I Poveda, Philip N Brown, Jason R Marden, and Andrew R Teel. "A class of distributed adaptive pricing mechanisms for societal systems with limited information". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2017, pp. 1490–1495.

[165] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. 1st. USA: John Wiley & Sons, Inc., 1994.

[166] Jairo Quiros-Tortos, Luis Ochoa, and Timothy Butler. "How electric vehicles and the grid work together: Lessons learned from one of the largest electric vehicle trials in the world". In: *IEEE Power and Energy Magazine* 16.6 (2018), pp. 64–76.

[167] Lillian J Ratliff and Tanner Fiez. "Adaptive incentive design". In: *IEEE Transactions on Automatic Control* 66.8 (2020), pp. 3871–3878.

[168] J Ben Rosen. "Existence and uniqueness of equilibrium points for concave n-person games". In: *Econometrica: Journal of the Econometric Society* (1965), pp. 520–534.

[169] Tim Roughgarden. "Algorithmic game theory". In: *Communications of the ACM* 53.7 (2010), pp. 78–86.

[170] William H Sandholm. *Population games and evolutionary dynamics*. MIT press, 2010.

[171] Shankar Sastry. *Nonlinear systems: analysis, stability, and control*. Vol. 10. Springer Science & Business Media, 2013.

[172] Kathrin Schacke. "On the kronecker product". In: *Master's thesis, University of Waterloo* (2004).

[173] AE Scheidegger. "Stability of motion, by nn krasovskii. translated from the (1959) russian ed. by jl brenner. stanford university press, 1963." In: *Canadian Mathematical Bulletin* 7.1 (1964), pp. 151–151.

[174] Mehran Shakarami, Ashish Cherukuri, and Nima Monshizadeh. "Dynamic interventions with limited knowledge in network games". In: *arXiv preprint arXiv:2205.15673* (2022).

[175] Mehran Shakarami, Ashish Cherukuri, and Nima Monshizadeh. "Steering the aggregative behavior of noncooperative agents: a nudge framework". In: *Automatica* 136 (2022), p. 110003.

[176] Kenneth A Small. "Using the revenues from congestion pricing". In: *Transportation* 19 (1992), pp. 359–381.

[177] David Roger Smart. *Fixed point theorems*. Vol. 66. Cup Archive, 1980.

[178] MJ Smith. "The marginal cost taxation of a transportation network". In: *Transportation Research Part B: Methodological* 13.3 (1979), pp. 237–242.

[179] Ziqi Song, Yafeng Yin, and Siriphong Lawphongpanich. "Nonnegative Pareto-improving tolls with multiclass network equilibria". In: *Transportation Research Record* 2091.1 (2009), pp. 70–78.

[180] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: An introduction.* MIT press, 2018.

[181] Brian Swenson, Ryan Murray, and Soummya Kar. "On best-response dynamics in potential games". In: *SIAM Journal on Control and Optimization* 56.4 (2018), pp. 2734–2767.

[182] Csaba Szepesvári. *Algorithms for reinforcement learning.* Synthesis lectures on artificial intelligence and machine learning. Morgan & Claypool, 2010.

[183] Tesla. *Introducing V3 Supercharging.* `https://www.tesla.com/blog/introducing-v3-supercharging`. 2019.

[184] Thomas C Thomas and Gordon I Thompson. "The value of time for commuting motorists as a function of their income level and amount of time saved". In: *Highway Research Record* 314 (1970).

[185] Pravin Varaiya. "Freeway Performance Measurement System: Final Report". In: *Univ. of California Berkeley, Berkeley, CA, USA, Tech. Rep. UCB-ITSPWP-2001-1* (2001).

[186] Erik T Verhoef. "Second-best congestion pricing in general networks. Heuristic algorithms for finding second-best optimal toll levels and toll points". In: *Transportation Research Part B: Methodological* 36.8 (2002), pp. 707–729.

[187] Roman Vershynin. *High-Dimensional Probability.* Cambridge University Press, 2018. ISBN: 1108415199.

[188] William G Waters. "The value of travel time savings and the link with income: implications for public project evaluation". In: *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti* (1994), pp. 243–253.

[189] Salomón Wollenstein-Betech, Chuangchuang Sun, Jing Zhang, and Ioannis Ch Paschalidis. "Joint estimation of od demands and cost functions in transportation networks from data". In: *2019 IEEE 58th Conference on Decision and Control (CDC).* IEEE. 2019, pp. 5113–5118.

[190] Di Wu, Yafeng Yin, Siriphong Lawphongpanich, and Hai Yang. "Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks". In: *Transportation Research Part B: Methodological* 46.9 (2012), pp. 1273–1287.

[191] Hai Yang and Hai-Jun Huang. *Mathematical and economic theory of road pricing.* Emerald Group Publishing Limited, 2005.

[192] Hai Yang and William HK Lam. "Optimal road tolls under conditions of queueing and congestion". In: *Transportation Research Part A: Policy and Practice* 30.5 (1996), pp. 319–332.

[193] Wei Yuan, Jianwei Huang, and Ying Jun Angela Zhang. “Competitive charging station pricing for plug-in electric vehicles”. In: *IEEE Transactions on Smart Grid* 8.2 (2015), pp. 627–639.

[194] Brian Hu Zhang, Gabriele Farina, Ioannis Anagnostides, Federico Cacciamani, Stephen Marcus McAleer, Andreas Alexander Haupt, Andrea Celli, Nicola Gatti, Vincent Conitzer, and Tuomas Sandholm. “Steering No-Regret Learners to Optimal Equilibria”. In: *arXiv preprint arXiv:2306.05221* (2023).

[195] Jing Zhang, Sepideh Pourazarm, Christos G Cassandras, and Ioannis Ch Paschalidis. “The price of anarchy in transportation networks by estimating user cost functions from actual traffic data”. In: *2016 IEEE 55th conference on decision and control (cdc)*. IEEE. 2016, pp. 789–794.

[196] Kai Zhang and Stuart Batterman. “Air pollution and health risks due to vehicle traffic”. In: *Science of the total Environment* 450 (2013), pp. 307–316.

[197] Lin Zhang, Haining Du, and Linda Lee. “Congestion Pricing Study of San Francisco–Oakland Bay Bridge in California”. In: *Transportation research record* 2221.1 (2011), pp. 83–95.

[198] Dao-Li Zhu, Hai Yang, Chang-Min Li, and Xiao-Lei Wang. “Properties of the multiclass traffic network equilibria under a tradable credit scheme”. In: *Transportation Science* 49.3 (2015), pp. 519–534.

Appendix A

Proofs for Chapter 2

Counter-example.

In this section, we present an instance of a non-atomic game where a standard gradient-based incentive design approach would fail. Specifically, we will show that the social cost, as a function of the incentive (evaluated at Nash equilibrium), is non-convex and has a zero gradient, which would result in the gradient-based approach not converging.

Consider a non-atomic routing game, comprising of two nodes and two edges connecting them. This network is used by one unit of travelers traveling from the source node $\mathcal{S}$ to the destination node $\mathcal{D}$. The latency function of two edges are denoted in Figure A.1. In this

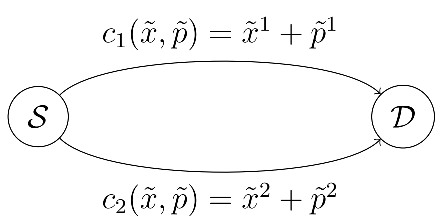


Figure A.1: Two-link routing game.

game, the strategy set is $\tilde{X} = \{\tilde{x} \in \mathbb{R}^2 : \tilde{x}_1 + \tilde{x}_2 = 1\}$. The equilibrium congestion levels on the two edges is obtained by computing the minimizer of the following function [191]:

$$T(\tilde{x},\tilde{p}) = \frac{1}{2}\tilde{x}_1^2 + \frac{1}{2}\tilde{x}_2^2 + \tilde{p}_1\tilde{x}_1 + \tilde{p}_2\tilde{x}_2.$$

Thus, for any toll vector $\tilde{p}$ the Nash equilibrium is $\tilde{x}^*(\tilde{p}) = \arg\min_{\tilde{x}\in\tilde{X}} T(\tilde{x},\tilde{p})$ is such that $\tilde{x}_1^*(\tilde{p}) = \mathcal{P}_{[0,1]}\left(\frac{\tilde{p}_2-\tilde{p}_1+1}{2}\right), \tilde{x}_2^*(\tilde{p}) = \mathcal{P}_{[0,1]}\left(\frac{\tilde{p}_1-\tilde{p}_2+1}{2}\right)$, where for any scalar $x \in \mathbb{R}$, $\mathcal{P}_{[0,1]}(x)$ denotes its projection on the line segment $[0,1]$.

The social cost function (compare (2.36)) is given by $\tilde{\Phi}(\tilde{x}) = \tilde{x}_1\tilde{\ell}_1(\tilde{x}) + \tilde{x}_2\tilde{\ell}_2(\tilde{x}) = \tilde{x}_1^2 + \tilde{x}_2^2$. Thus,

$$\tilde{\Phi}(\tilde{x}^*(\tilde{p})) = \begin{cases} \frac{(\tilde{p}_1-\tilde{p}_2)^2+1}{2} & \text{if } |\tilde{p}_1 - \tilde{p}_2| \leq 1, \\ 1 & \text{otherwise} \end{cases}$$

However, this function is not convex, for example, on the set $|\tilde{p}_1 - \tilde{p}_2| < 1.5$. Therefore, gradient based updates initialized in the region $\{(\tilde{p}_1, \tilde{p}_2) \in \mathbb{R}^2 : 1.1 \leq |\tilde{p}_1 - \tilde{p}_2| \leq 1.5\}$ will not converge to the global optimizer as the gradient is zero in this region. Thus, the approach from [121, 126] does not apply here.

Auxiliary Results

Lemma A.0.1 ([65]). *For any fixed p' and continuously differentiable function $\phi : \mathbb{R}^e \to \mathbb{R}^e$ the condition*

$$\langle \phi(p) - \phi(p'), p - p' \rangle \leq 0 \quad \forall\ p \in \mathcal{B}_r(p')$$

for some $r > 0$, holds if and only if $z^\top D\phi(p')z \leq 0$ for all $z \in \mathbb{R}^{|e|}$.

Lemma A.0.2. *The externality $\tilde{e}_i^j$ in (2.37) can be alternatively written, in terms of edge flows, as $\tilde{e}_i^j(\tilde{x}) = \sum_{a\in j} \tilde{w}_a \nabla l_a(\tilde{w}_a)$.*

Proof. Expanding the externality expression from (2.37) we observe that

$$\begin{aligned}
\tilde{e}_i^j(\tilde{x}) &= \sum_{i'\in\tilde{\mathcal{I}}}\sum_{j'\in\tilde{R}_i} \tilde{x}_{i'}^{j'} \frac{\partial \tilde{\ell}_{i'}^{j'}(\tilde{x})}{\partial \tilde{x}_i^j} \\
&\overset{(a)}{=} \sum_{i'\in\tilde{\mathcal{I}}}\sum_{j'\in\tilde{R}_i} \tilde{x}_{i'}^{j'} \sum_{a\in e} \mathbb{1}(a\in j') \frac{\partial l_a(\tilde{w}_a)}{\partial \tilde{x}_i^j} \\
&\overset{(b)}{=} \sum_{i'\in\tilde{\mathcal{I}}}\sum_{j'\in\tilde{R}_i} \tilde{x}_{i'}^{j'} \sum_{a\in e} \mathbb{1}(a\in j') \nabla l_a(\tilde{w}_a) \frac{\partial \tilde{w}_a}{\partial \tilde{x}_i^j} \\
&\overset{(c)}{=} \sum_{i'\in\tilde{\mathcal{I}}}\sum_{j'\in\tilde{R}_i} \tilde{x}_{i'}^{j'} \sum_{a\in e} \mathbb{1}(a\in j') \nabla l_a(\tilde{w}_a) \mathbb{1}(a\in j) \\
&\overset{(d)}{=} \sum_{a\in e} \nabla l_a(\tilde{w}_a) \mathbb{1}(a\in j) \sum_{i'\in\tilde{\mathcal{I}}}\sum_{j'\in\tilde{R}_i} \tilde{x}_{i'}^{j'} \mathbb{1}(a\in j') \\
&\overset{(e)}{=} \sum_{a\in j} \nabla l_a(\tilde{w}_a) \tilde{w}_a
\end{aligned}$$

where (a) follows by expanding out the expression of route costs in terms of edge costs, (b) follows by the chain rule, (c) follows due to the definition of edge flows, (d) follows by change of order of summations and finally (e) follows by the definition of edge flows. This completes the proof. □

Lemma A.0.3. *$x^\dagger$ satisfies* (2.42) *if and only if* (2.43).

Proof. First, we show that

$$\Phi(\tilde{x}) = \sum_{a\in e} \tilde{w}_a l_a(\tilde{w}_a). \tag{A.1}$$

Indeed,

$$\begin{aligned}
\Phi(\tilde{x}) &\overset{(2.36)}{=} \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \tilde{x}_i^j \tilde{\ell}_i^j(\tilde{x}) \\
&= \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \tilde{x}_i^j \sum_{a\in e} \mathbb{1}(a\in j) l_a(\tilde{w}_a) \\
&= \sum_{a\in e} l_a(\tilde{w}_a) \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \tilde{x}_i^j \mathbb{1}(a\in j) = \sum_{a\in e} l_a(\tilde{w}_a)\tilde{w}_a.
\end{aligned}$$

Next, observe that

$$\begin{aligned}
&\sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \frac{\partial\Phi(\tilde{x}^\dagger)}{\partial\tilde{x}_i^j}(\tilde{x}_i^j - \tilde{x}_i^{\dagger j}) \\
&= \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \sum_{a\in e} \frac{\partial}{\partial\tilde{x}_i^j}\left(\tilde{w}_a l_a(\tilde{w}_a)\right)(\tilde{x}_i^j - \tilde{x}_i^{\dagger j}) \\
&= \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \sum_{a\in e} \frac{\partial}{\partial\tilde{w}_a}\left(\tilde{w}_a l_a(\tilde{w}_a)\right)\mathbb{1}(a\in j)(\tilde{x}_i^j - \tilde{x}_i^{\dagger j}) \\
&= \sum_{a\in e} \frac{\partial}{\partial\tilde{w}_a}\left(\tilde{w}_a l_a(\tilde{w}_a)\right) \sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} \mathbb{1}(a\in j)(\tilde{x}_i^j - \tilde{x}_i^{\dagger j}) \\
&= \sum_{a\in e} \frac{\partial}{\partial\tilde{w}_a}\left(\tilde{w}_a l_a(\tilde{w}_a)\right)(\tilde{w}_a - \tilde{w}_a^\dagger).
\end{aligned}$$

This concludes the proof. □

Lemma A.0.4. *For any $\tilde{p}, \tilde{p}' \in \mathbb{R}^{|e|}$ the following holds*

$$\sum_{a\in e} (\tilde{p}_a - \tilde{p}'_a)(\tilde{w}_a^*(\tilde{p}_a) - \tilde{w}_a^*(\tilde{p}'_a)) \le 0. \tag{A.2}$$

Proof. To prove this result, we first show that

$$\sum_{i\in\tilde{\mathcal{I}}} \sum_{j\in\tilde{R}_i} (\tilde{P}_i^j - \tilde{P}_i^{\prime j})(\tilde{x}_i^{*j}(\tilde{P}) - \tilde{x}_i^{*j}(\tilde{P}')) \le 0, \tag{A.3}$$

where $\tilde{P}$ and $\tilde{P}'$ are the route tolls associated with edge tolls $\tilde{p}$ and $\tilde{p}'$ respectively through (2.35).

Let the feasible set of the optimization problem (2.39) be denoted by $\mathcal{F}$. Using the first-order conditions of optimality for the strictly convex optimization problem (2.39) we obtain the following variational inequality:

$$\sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}\left(\tilde{c}_i^j(\tilde{x}^*(\tilde{P}),\tilde{P})\right)\cdot\left(\tilde{y}_i^j-\tilde{x}_i^{*j}(\tilde{P})\right)\geq 0,\forall\quad \tilde{y}\in\mathcal{F}, \tag{A.4}$$

where $\tilde{P}$ is the route toll associated with edge toll $\tilde{p}$. Rewriting (A.4) for edge tolls $\tilde{p}'$ we obtain

$$\sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}\left(\tilde{c}_i^j(\tilde{x}^*(\tilde{P}'),\tilde{P}')\right)\cdot\left(\tilde{y}_i'^j-\tilde{x}_i^{*j}(\tilde{P}')\right)\geq 0,\forall\quad \tilde{y}'\in\mathcal{F}, \tag{A.5}$$

where $\tilde{P}'$ is the route toll associated with edge toll $\tilde{p}'$.

We are now ready to show the condition (A.3). Note that

$$\begin{aligned}
&\sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}(\tilde{P}_i^j-\tilde{P}_i'^j)(\tilde{x}_i^{*j}(\tilde{P})-\tilde{x}_i^{*j}(\tilde{P}'))\\
&\overset{(a)}{\leq}\sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}(\tilde{\ell}_i^j(\tilde{x}^*(\tilde{P}'))-\tilde{\ell}_i^j(\tilde{x}^*(\tilde{P})))(\tilde{x}_i^{*j}(\tilde{P})-\tilde{x}_i^{*j}(\tilde{P}'))\\
&\overset{(b)}{=}\sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}(\tilde{x}_i^{*j}(\tilde{P})-\tilde{x}_i^{*j}(\tilde{P}'))\cdot\\
&\qquad\qquad\cdot\sum_{a\in e}(l_a(\tilde{w}_a^*(\tilde{p}'))-l_a(\tilde{w}_a^*(\tilde{p})))\mathbb{1}(a\in j)\\
&\overset{(c)}{=}\sum_{a\in e}(l_a(\tilde{w}_a^*(\tilde{p}'))-l_a(\tilde{w}_a^*(\tilde{p})))\cdot\\
&\qquad\qquad\cdot\sum_{i\in\tilde{\mathcal{I}}}\sum_{j\in\tilde{R}_i}(\tilde{x}_i^{*j}(\tilde{P})-\tilde{x}_i^{*j}(\tilde{P}'))\mathbb{1}(a\in j)\\
&\overset{(d)}{=}\sum_{a\in e}(l_a(\tilde{w}_a^*(\tilde{p}'))-l_a(\tilde{w}_a^*(\tilde{p})))(\tilde{w}_a^*(\tilde{p})-\tilde{w}_a^*(\tilde{p}'))\\
&\overset{(e)}{\leq} 0
\end{aligned} \tag{A.6}$$

where we obtain (a) by adding (A.4), evaluated at $\tilde{y}=\tilde{x}^*(\tilde{P}')$, and (A.5), evaluated at $\tilde{y}'=\tilde{x}^*(\tilde{P})$, (b) holds by the definition of the route loss function, (c) holds by interchange of summation, (d) holds by the definition of edge flows and (e) holds due to the monotonicity of edge latency functions. This proves (A.3).

Now we are ready to prove (A.2). Note that

$$\sum_{a\in e}(\tilde{p}_a-\tilde{p}_{a'})(\tilde{w}_a^*(\tilde{p})-\tilde{w}_a^*(\tilde{p}'))$$

$$
\begin{aligned}
&\overset{(a)}{=} \sum_{a \in e} (\tilde{p}_a - \tilde{p}_{a'}) \sum_{i \in \tilde{\mathcal{I}}} \sum_{j \in \tilde{R}_i} (\tilde{x}_i^{*j}(\tilde{P}) - \tilde{x}_i^{*j}(\tilde{P}')) \mathbb{1}(a \in j) \\
&\overset{(b)}{=} \sum_{i \in \tilde{\mathcal{I}}} \sum_{j \in \tilde{R}_i} (\tilde{x}_i^{*j}(\tilde{P}) - \tilde{x}_i^{*j}(\tilde{P}')) \sum_{a \in e} (\tilde{p}_a - \tilde{p}_{a'}) \mathbb{1}(a \in j) \\
&\overset{(c)}{=} \sum_{i \in \tilde{\mathcal{I}}} \sum_{j \in \tilde{R}_i} (\tilde{x}_i^{*j}(\tilde{P}) - \tilde{x}_i^{*j}(\tilde{P}')) (\tilde{P}_i^j - \tilde{P}_i'^j) \\
&\overset{(d)}{\leq} 0,
\end{aligned}
$$

where (a) holds due to the definition of edge flows, (b) holds due to interchange of summation, (c) holds due to the definition of route tolls, (d) holds due to (A.3). This concludes the proof. □

Appendix B

Proofs for Chapter 3 and Additional Experiments

Preliminaries

Proof. (**Proof of Proposition 3.3.2**) Fix $x \in \mathbb{R}^n$ arbitrarily. By Bayes' rule:

$$\begin{aligned}
&\mathbb{P}(A_t = 1 | X_{t-1} \leq x, A_{1:t-1} = 0) \\
=&\mathbb{P}(A_t = 1 | A_{1:t-1} = 0) \\
&\quad \cdot \frac{\mathbb{P}(X_{t-1} \preceq x | A_t = 1, A_{1:t-1} = 0)}{\mathbb{P}(X_{t-1} \preceq x | A_{1:t-1} = 0)} \\
=&\mathbb{P}(A_t = 1 | A_{1:t-1} = 0) \cdot \alpha(x),
\end{aligned}$$

where $\alpha(x)$ is as defined in (3.2). Thus, $\mathbb{P}(A_t = 1 | A_{t-1} = 0)$ is time-invariant, i.e., holds the same value for each $t \in [T]$. Next, by invoking Bayes' rule again, we have:

$$\begin{aligned}
&\mathbb{P}(X_{T-1} \preceq x) \\
=&\sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x | T = t) \cdot \mathbb{P}(T = t) \\
=&\sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, T = t) \\
=&\sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0, A_t = 1) \\
=&\sum_{t=1}^{\infty} \mathbb{P}(A_t = 1 | X_t \preceq x, A_{1:t-1} = 0) \\
&\quad \cdot \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \\
=&a_1(x) \cdot \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0).
\end{aligned} \tag{B.1}$$

and:

$$\begin{aligned}&\sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \\ &\qquad \cdot \mathbb{P}(A_t = 1 | A_{1:t-1} = 0) \\ =& a_2 \cdot \sum_{t=1}^{\infty} \mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0).\end{aligned} \tag{B.2}$$

Thus, the null hypothesis H_0 in Definition 3.3.1 holds if and only if (B.1) and (B.2) are equal, as claimed. □

Methods

Proof. (**Proof of Theorem 3.4.1**) Fix $\epsilon > 0$, and take:

$$T_c := \left\lceil \frac{1}{\ln(1-p_1)} \ln\left(\frac{\epsilon p_1^2}{16 p_2}\right)\right\rceil.$$

First, to show that $\hat{b}_1^N(x) \to b_1(x)$ at an exponential rate in N, we invoke the Dvoretsky-Kiefer-Wolfowitz inequality:

$$\mathbb{P}\left(\sup_{x \in \mathbb{R}^n} \left|\hat{b}_1^N(x) - b_1(x)\right| > \frac{1}{2}\epsilon\right) \leq 2 \cdot e^{-\frac{1}{2}N\epsilon^2}$$

Next, to show that $\hat{b}_2^N(x) \to b_2(x)$ at an exponential rate in N, we have, via the triangle inequality:

$$\begin{aligned}&\sup_{x \in \mathbb{R}^n} \left|\hat{b}_2^N(x) - b_2(x)\right| \\ =& \sup_{x \in \mathbb{R}^n} \left|\sum_{t=1}^{\infty} \left[\hat{\beta}_t^N(x)\hat{\gamma}_t^N - \beta_t(x)\gamma_t\right]\right| \\ =& \sum_{t=1}^{T_c} \Big[\sup_{x \in \mathbb{R}^n} \left\{|\hat{\beta}_t^N(x) - \beta_t(x)|\right\}\hat{\gamma}_t^N \\ &\qquad + \sup_{x \in \mathbb{R}^n} \left\{|\hat{\gamma}_t^N - \gamma_t|\right\}\beta_t(x)\Big] \\ &\qquad + \sup_{x \in \mathbb{R}^n} \left\{\sum_{t=T_c+1}^{\infty} \left[|\hat{\beta}_t^N(x)\hat{\gamma}_t^N| + |\beta_t(x)\gamma_t|\right]\right\} \\ \leq& \sum_{t=1}^{T_c} \sup_{x \in \mathbb{R}^n} \left\{|\hat{\beta}_t^N(x) - \beta_t(x)|\right\}\end{aligned}$$

$$+\sum_{t=1}^{T_c} \sup_{x\in\mathbb{R}^n} \left\{|\hat{\gamma}_t^N - \gamma_t|\right\} \cdot \mathbb{P}(A_{1:t-1} = 0)$$
$$+\frac{1}{N}\sum_{i=1}^{N}\sum_{t=T_c+1}^{\infty} \mathbf{1}\{\hat{T}^i = t\} + \sum_{t=T_c+1}^{\infty} \mathbb{P}(T = t),$$

where the third and fourth term in the final expression follow by observing that, for any $x \in \mathbb{R}$, by definition of the quantities $\hat{\beta}_t^N(x)$, $\hat{\gamma}_t^N$, $\beta_t(x)$, and γ_t:

$$\begin{aligned}
&|\hat{\beta}_t^N(x)\hat{\gamma}_t^N| \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{X_t^i \preceq x, A_{1:t-1}^i = 0\} \\
&\qquad \cdot \frac{\sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0, A_t^i = 1\}}{\sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0\}} \\
\leq&\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{A_{1:t-1}^i = 0\} \\
&\qquad \cdot \frac{\sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0, A_t^i = 1\}}{\sum_{i=1}^N \mathbf{1}\{A_{1:t-1}^i = 0\}} \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{A_{1:t-1}^i = 0, A_t^i = 1\} \\
=&\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{\hat{T}^i = t\}
\end{aligned}$$

and similarly:

$$\begin{aligned}
&|\beta_t(x)\gamma(t)| \\
=&\mathbb{P}(X_{t-1} \preceq x, A_{1:t-1} = 0) \cdot \mathbb{P}(A_t = 1 | A_{1:t-1} = 0) \\
\leq&\mathbb{P}(A_{1:t-1} = 0) \cdot \mathbb{P}(A_t = 1 | A_{1:t-1} = 0) \\
\leq&\mathbb{P}(A_t = 1, A_{1:t-1} = 0) \\
=&\mathbb{P}(T = t).
\end{aligned}$$

Below, we upper bound each of the four terms in the final expression above.

- First, by the Dvoretsky-Kiefer-Wolfowitz inequality, we have, for each $t \in [T_c] := \{1, \cdots, T_c\}$:

$$\mathbb{P}\left(\sum_{t=1}^{T_c} \sup_{x\in\mathbb{R}^n} \left\{|\hat{\beta}_t^N(x) - \beta_t(x)|\right\} \geq \frac{1}{8}\epsilon\right)$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\bigcup_{t=1}^{T_c}\left\{\sup_{x\in\mathbb{R}^n}\left\{|\hat{\beta}_t^N(x)-\beta_t(x)|\right\}\ge \frac{1}{8T_c}\epsilon\right\}\right)\\
&\le \sum_{t=1}^{T_c}\mathbb{P}\left(\sup_{x\in\mathbb{R}^n}\left\{|\hat{\beta}_t^N(x)-\beta_t(x)|\right\}\ge \frac{1}{8T_c}\epsilon\right)\\
&\le 2T_c\exp\left(-\frac{\epsilon^2}{32T_c^2}\cdot N\right).
\end{aligned}
$$

- Second, let $N_t \in [N]$ denote the number of trajectories with $A_{1:t-1}=0$. We first show that, with high probability, $N_t \ge N\cdot\mathbb{P}(A_{1:t-1}=0)^2$. We then show that, under this condition on N_t taking a sufficiently large value, $\hat{\gamma}_t^N(x)\rangle\gamma_t(x)$ exponentially in N.

  First, the Hoeffding bound for general bounded random variables ([187] Theorem 2.2.6) gives:

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{N}N_t \le \mathbb{P}(A_{1:t-1}=0)^2\right)\\
\le&\mathbb{P}\Bigg(\left|\frac{1}{N}N_t-\mathbb{P}(A_{1:t-1}=0)\right|\\
&\qquad\ge \mathbb{P}(A_{1:t-1}=0)-\mathbb{P}(A_{1:t-1}=0)^2\Bigg)\\
\le&\exp\left(-2\Big[\mathbb{P}(A_{1:t-1}=0)-\mathbb{P}(A_{1:t-1}=0)^2\Big]^2N\right)
\end{aligned}
$$

  Then, if $N_t \ge N\cdot\mathbb{P}(A_{1:t-1}=0)$, we can bound the gap between $\hat{\gamma}_t^N(x)$ and $\gamma_t(x)$ as follows:

$$
\begin{aligned}
&\mathbb{P}\left(|\hat{\gamma}_t^N(x)-\gamma_t(x)|>\frac{\epsilon}{8T_c\cdot\mathbb{P}(A_{1:t-1}=0)}\right)\\
&\le\exp\Bigg(-2\cdot\mathbb{P}(A_{1:t-1}=0)^2\cdot N\\
&\qquad\cdot\frac{\epsilon^2}{64T_c^2\cdot\mathbb{P}(A_{1:t-1}=0)^2}\Bigg)\\
&\le\exp\left(-\frac{\epsilon^2}{32T_c^2}\cdot N\right).
\end{aligned}
$$

- Third, to bound $\hat{B}_{T_c}^N := \frac{1}{N}\sum_{i=1}^N\sum_{t=T_c+1}^{\infty}\mathbf{1}\{\hat{T}^i=t\}$, define:

$$
B_{T_c} := \sum_{t=T_c+1}^{\infty}\mathbf{1}\{T=t\}=\mathbf{1}\{T>T_c\}.
$$

Thus, B_{T_c} is a Bernoulli random variable with parameter $\mathbb{P}(B_{T_c} = 1)$, and expectation upper bounded by:

$$\mathbb{E}[B_{T_c}] = \mathbb{P}(B_{T_c} = 1) \leq (1 - p_1)^{T_c}.$$

By definition of T_c, we have $\mathbb{E}[B_{T_c}] \leq \frac{1}{16}\epsilon$. Moreover, since B_{T_c} is a Bernoulli random variable, we have, by the Hoeffding bound for general bounded random variables ([187] Theorem 2.2.6):

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{B}_{T_c}^{N} > \frac{1}{8}\epsilon\right) \\
=&\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{B}_{T_c}^{N} - \mathbb{E}[B_{T_c}] > \frac{1}{8}\epsilon - \mathbb{E}[B_{T_c}]\right) \\
\leq&\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{B}_{T_c}^{N} - \mathbb{E}[B_{T_c}] > \frac{1}{16}\epsilon\right) \\
<& \exp\left(-\frac{1}{128}N\right).
\end{aligned}$$

- Finally, note that by definition of T_c:

$$\begin{aligned}
\sum_{t=T_c+1}^{\infty}\mathbb{P}(T = t) &= \mathbb{P}(T > T_c) \\
&\leq (1 - p_1)^{T_c} \\
&< \frac{1}{16}\epsilon.
\end{aligned}$$

For the multivariate version (i.e., $n > 1$), the same proof follows, albeit with the multivariate version of the Dvoretsky-Kiefer-Wolfowitz inequality [144]. □

Experiment Results

Multi-link Traffic Networks

For the multi-link traffic network, we use the dynamics: ([130])

$$x_i[t+1] \tag{B.3}$$

$$= (1-\mu)\cdot x_i[t] + \mu \cdot \frac{e^{-\beta\cdot x_i[t]}}{\sum_{j=1}^{R} e^{-\beta\cdot x_j[t]}} \cdot u[t] + w[t], \tag{B.4}$$

$$\forall t \in [T], i \in [R], \tag{B.5}$$

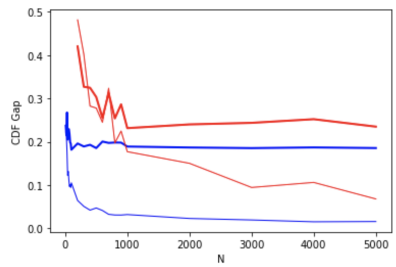


Figure B.1: CDF Gap between vs. N, for the 2-link traffic network example. Here, red and blue correspond to the baseline and our method, respectively, while thick and thin lines correspond to the null and alternative hypotheses, respectively. Our approach correctly identifies the null hypothesis dataset with a relatively small number of samples, while the naive aggregation method fails to do so (thin blue curve).

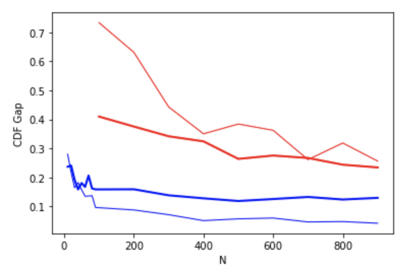


Figure B.2: CDF Gap between vs. N, for the 3-link traffic network example. The color and thickness schemes are identical to those of the single-link and 2-link plots in Figures 3.2 and B.1.

$$A[t] \sim \mathcal{P}(x[t]), \tag{B.6}$$

where $x_i[t]$ denotes the traffic flow on each link $i \in [R]$, $u[t] \in \mathbb{R}$ and $w[t] \in \mathbb{R}$, and T_h, are the input, zero-mean noise terms, and time horizon, as before. Here, we set $T = 250$, $\mu(0) = 0.3$, $\mu(1) = 0.2$, $u(t) = 100R$ for each $t \in [T]$, and we again draw $w[t]$ i.i.d. from the continuous uniform distribution on $(-10, 10)$. As with the single-link case, we created two datasets for the null and alternative hypotheses. For the null hypothesis, we fix $\mathcal{P}(x[t])$ to be Bernoulli(0.02); for the alternative hypothesis, we set $\mathcal{P}(x[t])$ to be Bernoulli(0.02) when $x[t] < 105$, and Bernoulli(0.30) when $x[t] \geq 105$. Again, this setting encodes the situation where higher traffic loads cause higher accident probabilities.

Similar to the single-link case, we compute the maximum CDF gap $\sup_{x \in \mathbb{R}^n} |\hat{b}_1^N(x) - \hat{b}_2^N(x)|$ as functions of N (thin lines), and the empirical CDFs of $X_{t-1}|T = t$ and X_{t-1} (thick lines) for both the null and alternative hypotheses. We again observe that our method distinguishes between the two hypotheses at a smaller sample number N compared to the baseline method.

Analogous results hold for a 3-link system with dynamics as given by (B.3) and are presented in Figure B.2.

Appendix C

Proofs for Chapter 4

Proof of Lemma 4.4.2

Note that $\Theta = I_T \otimes \big(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + \mathbf{C}) + \mu\big)$ is positive definite if and only if each of the terms in the Kronecker product is positive definite [172]. Hence, it is sufficient to show that $Q := b(\mathbf{1}_N\mathbf{1}_N^\top + \mathbf{C}) + \mu$ is a positive definite matrix. First, note that Q is a symmetric matrix. Additionally, for any $z \in \mathbb{R}^N$, it holds that

$$z^\top Q z = b(\mathbf{1}_N^\top z)^2 + b(\sum_{i\in\mathcal{C}} z_i)^2 + \sum_{i\in\mathcal{C}} \mu_i z_i^2 + \sum_{i\in N\setminus\mathcal{C}} (b+\mu_i) z_i^2,$$

which is strictly positive for all $z \neq 0$. Next, we show that Γ is invertible by calculating its inverse. Indeed,

$$\begin{aligned}\Gamma^{-1} &= \big((\mathbf{1}_T^\top \otimes I_\mathcal{I})\Theta^{-1}(\mathbf{1}_T \otimes I_\mathcal{I})\big)^{-1} \\ &= \Big(\mathbf{1}_T^\top\mathbf{1}_T \otimes \big(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + C) + \mu\big)^{-1}\Big)^{-1} \\ &= \frac{1}{T}\big(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + C) + \mu\big) = \frac{1}{T} Q.\end{aligned}$$

This completes the proof.

Proof of Theorem 4.4.6

For every $t \in [T]$, define $G^t := \sum_{t'\in[T]} a^{t'}\big(T\delta^{tt'} - 1\big)$. Using this, the expression for $x^{\dagger t}$ from Proposition 9 can be re-written as follows

$$x^{\dagger t} = \frac{d}{T} - \frac{1}{T}\big(b(\mathbf{1}_\mathcal{I}\mathbf{1}_\mathcal{I}^\top + C) + \mu\big)^{-1}\mathbf{1}_\mathcal{I} G^t. \tag{C.1}$$

Consequently, it holds that

$$\begin{aligned} x_i^{\dagger t} &= \frac{d_i}{T} - \frac{G^t}{T}\Gamma_i^\dagger \\ \mathbf{1}_N^\top x^{\dagger t} &= \frac{D}{T} - \frac{G^t}{T}\Gamma^\dagger \\ x_i^{\dagger t} - \bar{x}_i^t &= -\frac{1}{T}\Gamma_i^\dagger G^t. \end{aligned} \tag{C.2}$$

Consequently, combining previous equations with (4.1), for every $i \in [N]$,

$$\begin{aligned} c_i(x^\dagger) &= \sum_{t\in[T]} a^t x_i^{\dagger t} + \frac{b}{T^2}\sum_{t\in[T]} \left(D - G^t\Gamma^\dagger\right)\left(d_i - G^t\Gamma_i^\dagger\right) \\ &\quad + \frac{\mu_i(\Gamma_i^\dagger)^2}{2T^2}\sum_{t\in[T]}(G^t)^2 \\ &= \sum_{t\in[T]} a^t \left(\frac{d_i}{T} - \frac{G^t}{T}\Gamma_i^\dagger\right) + \frac{b}{T^2}\left(TDd_i + \Gamma^\dagger\Gamma_i^\dagger \sum_{t\in[T]}(G^t)^2\right) \\ &\quad + \frac{\mu_i(\Gamma_i^\dagger)^2}{2T^2}\sum_{t\in[T]}(G^t)^2, \end{aligned} \tag{C.3}$$

where the last equality is because $\sum_{t\in[T]} G^t = \sum_{t,t'\in[T]} a^{t'}\left(T\delta^{tt'} - 1\right) = 0$. Additionally, summing (C.3) for all $i \in \mathcal{S}$, we obtain

$$\begin{aligned} \sum_{i\in\mathcal{S}} c_i(x^\dagger) &= \sum_{t\in[T]} a^t \left(\frac{D_\mathcal{S}}{T} - \frac{G^t}{T}\sum_{i\in\mathcal{S}}\Gamma_i^\dagger\right) + \frac{bDD_\mathcal{S}}{T} \\ &\quad + \frac{b\Gamma^\dagger \sum_{i\in\mathcal{S}}\Gamma_i^\dagger \sum_{t\in[T]}(G^t)^2}{T^2} + \frac{b}{2T^2}\sum_{t\in[T]}(G^t)^2 \sum_{i\in\mathcal{S}}\frac{\mu^i}{b}(\Gamma_i^\dagger)^2 \\ &= \left(\frac{AD_\mathcal{S} - \sum_{i\in\mathcal{S}}\Gamma_i^\dagger \sum_{t\in[T]} a^t G^t}{T} + \frac{bDD_\mathcal{S}}{T}\right) \\ &\qquad + \frac{b\sum_{t\in[T]}(G^t)^2}{T^2}\left(\Gamma^\dagger \sum_{i\in\mathcal{S}}\Gamma_i^\dagger + \sum_{i\in\mathcal{S}}\frac{\mu^i}{2b}(\Gamma_i^\dagger)^2\right), \end{aligned}$$

where $A := \sum_{t\in[T]} a^t$. Analogously, we obtain the cost at Nash equilibrium as follows

$$\begin{aligned} \sum_{i\in\mathcal{S}} c_i(x^*) &= \left(\frac{AD_\mathcal{S} - \sum_{i\in\mathcal{S}}\Gamma_i^* \sum_{t\in[T]} a^t G^t}{T} + \frac{bDD_\mathcal{S}}{T}\right) \\ &\qquad + \frac{b\sum_{t\in[T]}(G^t)^2}{T^2}\left(\Gamma^* \sum_{i\in\mathcal{S}}\Gamma_i^* + \sum_{i\in\mathcal{S}}\frac{\mu^i}{2b}(\Gamma_i^*)^2\right). \end{aligned}$$

Using the above equations, we obtain that $\sum_{i\in\mathcal{S}} c_i(x^\dagger) > \sum_{i\in\mathcal{S}} c_i(x^*)$ if and only if $g^\dagger_{\mathcal{S}}(\mu, b) - g^*_{\mathcal{S}}(\mu, b) \geq (\Gamma^\dagger - \Gamma^*)h(A)$. This completes the proof.

Appendix D

Proofs for Chapter 6

We begin by proving theorem 6.4.1, which adapts techniques from analogous results in the social choice literature (i.e. theorem 6.2.6):

Proof. First, we show that for any reward function R in a Social Choice MDP, f_R satisfies Universal Domain, Completeness and Transitivity. Consider an arbitrary profile $U \in \mathscr{U}$. We have $x f_R(U) y \iff R(U,x) \geq R(U,y)$, which is well defined since the domain of R is $\mathscr{U} \times X$. Therefore, f_R satisfies Universal Domain. Consider an arbitrary profile $U \in dom(f_R)$. By completeness of $\geq$ on $\mathbb{R}$, we have $R(U,x) \geq R(U,y)$ or $R(U,y) \geq R(U,x)$, so $x f_R(U) y$ or $y f_R(U) x$, so $f_R(U)$ is complete. Now assume $x f_R(U) y$ and $y f_R(U) z$ for some $x, y, z \in X$. It follows that $R(U,x) \geq R(U,y)$ and $R(U,y) \geq R(U,z)$. Therefore, $R(U,x) \geq R(U,z)$, so $x f_R(U) z$. Therefore, $f_R(U)$ is transitive. Since U was arbitrary, f_R satisfies Completeness and Transitivity.

Assume, in addition, that R satisfies **Weak Pareto**, **IIA**, **CUC-Invariance** and **Functional Anonymity**. Therefore, f_R satisfies **Pareto (SWF)**, **IIA**, **CUC-Invariance** and **Functional Anonymity**. So by theorem 6.2.6, f_R is Utilitarianism:

$$x f_R(U) y \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y).$$

By definition of f_R, we have $x f_R(U) y \iff R(U,x) \geq R(U,y)$, so

$$R(U,x) \geq R(U,y) \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y),$$

so R agrees with Utilitarianism.

We now prove the converse direction of the equivalence. That is, we want to show that if R agrees with Utilitarianism, that is, for any profile U and $x, y \in X$, we have

$$R(U,x) \geq R(U,y) \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y) \tag{D.1}$$

then R satisfies **Pareto (SWF)**, **IIA**, **CUC-Invariance**, and **Functional Anonymity**.

We start by showing that R satisfies **Pareto (SWF)**. That is, we want to show that for all $U \in dom(f_R)$ and all $x, y \in X$, if $U_i(x) > U_i(y)$ for all $i \in V$, then $xP(f_R(U))y$. Assume that $U_i(x) > U_i(y)$ for all $i \in V$. Then, we know that $\sum_{i \in V} U_i(x) > \sum_{i \in V} U_i(y)$. Because R agrees with Utilitarianism, then we know that $R(U, x) > R(U, y)$, and in turn, this means that the social welfare functional f_R induced by R satisfies $xP(f_R(U))y$ for all $x, y \in X$, which is what we wanted to show.

Next, we consider **IIA**. That is, we want to show that for all $U, U' \in dom(f_R)$ and $x, y \in X$, if $U_i(x) = U'_i(x)$ and $U_i(y) = U'_i(y)$ for all $i \in V$, then $xf_R(U)y$ if and only if $xf_R(U')y$. Assume that $U_i(x) = U'_i(x)$ and $U_i(y) = U'_i(y)$ and $xf_R(U)y$. Then, because f_R is the SWF induced by R, we know that $R(U, x) \geq R(U, y)$, and furthermore, because R agrees with Utilitarianism, we know that $\sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y)$. However, by the property that $U_i(x) = U'_i(x)$ and $U_i(y) = U'_i(y)$ for all $i \in V$, we get the inequality $\sum_{i \in V} U'_i(x) \geq \sum_{i \in V} U'_i(y)$. Therefore, once again, because R agrees with Utilitarianism, we have $R(U', x) \geq R(U', y)$, and thus $xf_R(U')y$. The steps in this proof are reversible, and thus the converse direction follows as well.

Now consider **CUC-Invariance**. For all $U, U' \in dom(f)$, we want to show that if $U \sim_{CUC} U'$, then $f_R(U) = f_R(U')$. Assume $U \sim_{CUC} U'$. By definition, there is a $\beta \in \mathbb{R}$ with $\beta > 0$ and for every $i \in V$, there is some $\alpha_i \in \mathbb{R}$ such that for all $x \in X$, $U_i(x) = \alpha_i + \beta U'_i(x)$. We have, for all $x, y \in X$,

$$xf_R(U)y \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y)$$

by assumption. By standard properties of summation,

$$\sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y) \iff \sum_{i \in V} \beta U_i(x) \geq \sum_{i \in V} \beta U_i(y) \iff \sum_{i \in V} \alpha_i + \beta U_i(x) \geq \sum_{i \in V} \alpha_i + \beta U_i(y)$$

and by definition

$$\sum_{i \in V} \alpha_i + \beta U_i(x) \geq \sum_{i \in V} \alpha_i + \beta U_i(y) \iff \sum_{i \in V} U'_i(x) \geq \sum_{i \in V} U'_i(y) \iff xf_R(U')y.$$

Therefore, $xf_R(U)y \iff xf_R(U')y$, so $f_R(U) = f_R(U')$.

We finish by showing **Functional Anonymity**. We want to show that for all $U, U' \in dom(f_R)$ and permutations $\rho : V \to V$, if for all $i \in V$, $U'_i = U_{\rho(i)}$, then $f_R(U) = f_R(U')$. Consider profiles U and U' and a permutation $\rho : V \to V$ such that for all $i \in V$, $U'_i = U_{\rho(i)}$. Now, for all $x, y \in X$:

$$xf(U)y \iff \sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y)$$

Permutations do not affect the sum, so we have, for any permutation $\rho : V \to V$,

$$\sum_{i \in V} U_i(x) \geq \sum_{i \in V} U_i(y) \iff \sum_{i \in V} U_{\rho(i)}(x) \geq \sum_{i \in V} U_{\rho(i)}(y),$$

and by definition

$$\sum_{i\in V} U_{\rho(i)}(x) \geq \sum_{i\in V} U_{\rho(i)}(y) \iff \sum_{i\in V} U'_i(x) \geq \sum_{i\in V} U'_i(y) \iff x f_R(U') y,$$

which completes our proof. □

We proceed by proving theorem 6.4.6:

Proof. Consider a social choice MDP where V satisfies the Bellman equation and R satisfies **Weak Pareto**, **IIA**, **CUC-Invariance** and **Functional Anonymity**.

Assume that π is optimal relative to V. Therefore, for all $s \in \mathcal{S}$

$$\pi \in \arg\max_{\pi\in\Pi} V(\pi, s).$$

Since V satisfies the Bellman equation, we have $V(\pi, s) = \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t R(s_t, \pi(s_t))\right]$ for all $s \in \mathcal{S}$ by theorem 6.4.4, so

$$\pi^* \in \arg\max_{\pi\in\Pi} \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t R(s_t, \pi(s_t))\right].$$

for all $s \in \mathcal{S}$.

By **Weak Pareto**, **IIA**, **CUC-Invariance** and **Functional Anonymity** and theorem 6.4.1, R is quasi-utilitarian, so we have $R(U, a) = f\left(\sum_{i\in V} U_i(a)\right)$, where $f : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function. Therefore,

$$\pi^* \in \arg\max_{\pi\in\Pi} \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t f\left(\sum_{i\in V} U_i^t(a)\right)\right],$$

where f is strictly increasing, so π is long-run quasi utilitarian.

Assume that π is long-run quasi utilitarian. By definition, for all $s \in \mathcal{S}$

$$\pi^* \in \arg\max_{\pi\in\Pi} \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t f\left(\sum_{i\in V} U_i^t(a)\right)\right],$$

so

$$\pi^* \in \arg\max_{\pi\in\Pi} \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t R(s_t, \pi(s_t))\right],$$

where $R(U, a) = f\left(\sum_{i\in V} U_i^t(a)\right)$ for some strictly increasing f. By theorem 6.4.4, since V satisfies the Bellman equation, we have $V(\pi, s) = \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t R(s_t, \pi(s_t))\right]$ for all $s \in \mathcal{S}$.

Therefore, we have for all $s \in \mathcal{S}$

$$\pi \in \arg\max_{\pi\in\Pi} V(\pi, s),$$

so π is optimal relative to V. □