Credit-Based Congestion Pricing: Equilibrium Properties and Optimal Scheme Design

Devansh Jalota[†], Jessica Lazarus[‡], Alexandre Bayen[*], and Marco Pavone[†]

September 2023

Abstract

Credit-based congestion pricing (CBCP) has emerged as a mechanism to alleviate the social inequity concerns of road congestion pricing - a promising strategy for traffic congestion mitigation - by providing low-income users with travel credits to offset some of their toll payments. While CBCP offers immense potential for addressing inequity issues that hamper the practical viability of congestion pricing, the deployment of CBCP in practice is nascent, and the potential efficacy and optimal design of CBCP schemes have yet to be formalized. In this work, we study the design of CBCP schemes to achieve particular societal objectives and investigate their influence on traffic patterns when routing heterogeneous users with different values of time (VoTs) in a multi-lane highway with an express lane. We introduce a new non-atomic congestion game model of a *mixed-economy*, wherein eligible users receive travel credits while the remaining ineligible users pay *out-of-pocket* to use the express lane. In this setting, we investigate the effect of CBCP schemes on traffic patterns by characterizing the properties (i.e., existence, comparative statics) of the corresponding Nash equilibria and, in the setting when eligible users have time-invariant VoTs, develop a convex program to compute these equilibria. We further present a bi-level optimization framework to design optimal CBCP schemes to achieve a central planner's societal objectives. Finally, we conduct numerical experiments based on a case study of the San Mateo 101 Express Lanes Project, one of the first North American CBCP pilots. Our results demonstrate the potential of CBCP to enable low-income travelers to avail of the travel time savings provided by congestion pricing on express lanes while having comparatively low impacts on the travel costs of other road users.

1 Introduction

With the ever-worsening traffic congestion in urban metropolises, congestion pricing has emerged as one of the most promising traffic management policies across the world [11, 12]. The premise behind congestion pricing is to charge users for the negative externality they impose on others to reduce system inefficiencies caused by selfish travel behavior [62, 63]. In effect, congestion pricing serves to steer the user equilibrium traffic pattern towards the system optimum traffic pattern [66, 15, 28]. While network-wide deployments of congestion pricing, wherein tolls are placed on all or some cordoned portion of roads in the road network, are less common at present, there has been a growing interest in introducing congestion fees on certain lanes on highways, known as *express lanes* or managed lanes, to provide users with a faster and more reliable travel option during peak traffic periods. In the Bay Area alone, there are more than 155 miles of express lanes at the time of writing this manuscript [18], with varying tolls depending on the region and the length of the highway section. In many cases, existing *high-occupancy vehicle* (HOV) lanes, which only grant access to vehicles with more than two or three passengers, have been converted to *high-occupancy toll (HOT)* lanes that enable single-occupant vehicles (SOVs) to pay for access.

Despite the proliferation of express lanes to better manage highway traffic, congestion fees on express lanes, as with network-wide congestion pricing schemes, have come under scrutiny due to social inequity concerns. In particular, express lanes have been termed as elitist "Lexus lanes", as they offer only those

---

[†] Stanford University; {djalota, pavone}@stanford.edu.
[‡] Lawerence Berkeley National Laboratory; jlazarus@lbl.gov.
[‡] University of California Berkeley; bayen@berkeley.edu.

with the highest willingness to pay (i.e., the most wealthy) a higher quality of service through reduced travel times [60, 58, 33] while lower-income users bear the brunt of longer travel times on consequently more congested general purpose (GP) lanes. More generally, there is a widespread belief that congestion pricing amounts to a "tax on the working class [59]," given its regressive nature [24, 51, 29]. Thus, there has been a growing interest in designing equitable congestion pricing mechanisms [72], with a focus on designing *credit-based congestion pricing* (CBCP) schemes [45]. In CBCP schemes, road users, particularly those with lower incomes, are given travel credits to use priced roads such as express lanes to counter the adverse equity impacts of congestion pricing. Although numerous variations of CBCP have been explored in the literature, demonstrating their potential to provide positive equity benefits, their deployment in practice is nascent, with one of the first North American pilots launching in San Mateo County, California, in 2022. In particular, the *San Mateo 101 Express Lanes Project*, which is converting 22 miles of HOV lanes to HOT express lanes, recently launched the "Community Transportation Benefits Program" to improve equitable access to the new facility [23]. The equity program provides low-income residents with a $100 credit for the express lane, enabling them to use the express lane for free, which they may not have been able to avail of otherwise, and thus reduce their travel times.

While CBCP programs such as the one deployed in San Mateo County offer great potential to improve equity outcomes, a principled design of CBCP schemes is necessary to realize the benefits of its implementation. To this end, we study CBCP schemes to route heterogeneous users with different values of time (VoTs) on a multi-lane highway segment with express lanes. In particular, we introduce a *mixed-economy* model, wherein eligible users receive travel credits to use the express lane while ineligible users have quasi-linear costs as they pay *out-of-pocket* to access the express lane. We note that our mixed-economy model is unlike traditional *single-economy* settings in traffic routing, wherein either all users have quasi-linear costs [74] or budget constraints [25] as in markets with artificial currencies. In this mixed-economy setting, we study the influence of CBCP schemes on the resulting traffic patterns by characterizing the properties of the induced equilibria and present a bi-level optimization framework to design optimal CBCP schemes. Furthermore, we demonstrate the real-world applicability of our proposed model with numerical experiments based on a case study of the San Mateo 101 Express Lanes Project.

1.1   Our Contributions

We study CBCP schemes in non-atomic congestion games to route heterogeneous users with varying VoTs on a multi-lane highway segment with a tolled express lane. In alignment with practically deployed CBCP schemes, such as in the San Mateo 101 Express Lanes Project, we introduce a new model of a mixed-economy wherein eligible users receive travel credits while ineligible users pay out-of-pocket to use the express lane. Given the different optimization objectives of eligible and ineligible users, we introduce a new Nash equilibrium notion, which we term CBCP equilibria, to investigate the influence of a CBCP scheme on the resulting traffic patterns.

In this mixed economy setting, we begin by investigating the properties of CBCP equilibria. In particular, we establish the existence of CBCP equilibria, prove the uniqueness of the resulting edge flows, and, in the setting when eligible users have time-invariant VoTs, develop a convex program to compute CBCP equilibria. We further perform a comparative statics analysis to investigate the changes in the equilibria induced by CBCP schemes when the express lane tolls or eligible user budgets are increased or decreased. In particular, we establish several natural monotonicity relations between the aggregate edge flows and the express lane tolls and budgets of eligible users that align with economic intuition. However, we do mention that we also obtain some counter-intuitive results, such as the violation of a natural substitutes condition (see Section 5.2), due to the introduction of travel credits for eligible users.

We then study the design of optimal CBCP schemes to achieve specific societal objectives of a central planner in the setting when eligible users have time-invariant VoTs. In this context, we develop a bi-level optimization framework for designing optimal CBCP schemes and present a *dense sampling* approach for computing an approximation to the optimal CBCP scheme. Our dense sampling approach involves discretizing the set of feasible CBCP schemes and choosing the scheme that induces an equilibrium with the minimum societal cost. We further motivate the efficacy of our dense sampling approach by establishing continuity relations between the resulting CBCP equilibria (i.e., the aggregate edge flows in the network) and the corresponding toll and budget parameters that characterize a CBCP scheme. In particular, our

obtained continuity relations imply that, for a small enough discretization, the CBCP scheme output by dense sampling is a good approximation to the optimal CBCP scheme, as small changes in the toll and budget parameters only result in small changes in the societal cost function.

Finally, we present numerical experiments to investigate the influence of CBCP schemes on traffic patterns and study their optimal design through a case study of the San Mateo 101 Express Lanes Project. Our numerical results validate our comparative statics analysis results and indicate that the magnitude of the changes in the travel time and proportion of users on the express lane with the tolls and budgets observed in the experiments are reflective of real-world multi-lane highways with express lanes. Furthermore, since the optimal CBCP scheme can vary widely based on the central planner's objective, our results show that a principled approach using bi-level optimization proposed in this work is key to realizing the benefits of CBCP schemes. We also discuss the policy implications of this work, noting avenues for future work and considerations for further mitigating the inequity concerns of congestion pricing.

Organization: This paper is organized as follows. Section 2 reviews related literature. We then present a model of traffic flow and introduce a new Nash equilibrium notion, the CBCP equilibrium, studied in this work in Section 3. Next, we investigate the existence of Nash equilibria induced by CBCP schemes in Section 4 and perform a comparative statics analysis of the CBCP equilibria in Section 5. Then, we introduce a bi-level optimization framework to design optimal CBCP schemes in Section 6 and present numerical experiments based on a real-world case study of the San Mateo 101 Express Lanes Project in Section 7. Finally, we conclude the paper and provide directions for future work in Section 8.

2 Related Literature

While achieving system efficiency in resource allocation is often the holy grail for central planners, there are many settings when a range of measures of user welfare beyond system efficiency, e.g., fairness or equity considerations, need to be taken into account to design practically deployable allocation mechanisms. For instance, in a seminal work, [71] showed that there are many settings, e.g., cancer treatment, when it is undesirable to efficiently allocate resources to users with the highest willingness to pay. In a similar spirit, [17] studied the allocation of identical objects to users and showed that the choice between market and non-market allocation mechanisms crucially relies on the similarity between a central planner's objective and users' willingness to pay. Furthermore, [3] developed a resource allocation model to allocate heterogeneous objects to a continuum of agents to trade off allocative efficiency with redistributive considerations. More generally, this trade-off between system efficiency and other welfare measures, such as fairness or equity, has been investigated in a range of resource allocation applications, including vaccine allocation [2], aircraft scheduling [7], and traffic routing [38, 41].

The consideration of welfare measures beyond system efficiency is, in particular, important in the context of congestion pricing, as, despite several successful real-world deployments [30, 53, 49], its practical adoption has been quite limited [24, 51, 37], in part due to social inequity concerns. Thus, there has been a growing interest in designing more equitable congestion pricing schemes [72], with a particular focus on revenue redistribution approaches [67], wherein the collected toll revenues are refunded as lump-sum transfers to users. In particular, several works [5, 20, 1, 32] consider the problem of refunding the collected toll revenues to users in problem settings that range from Vickrey's bottleneck congestion model [69] - a benchmark representation of peak-period traffic congestion on a single lane - to general road networks with a multiple origin-desitination travel demand and heterogeneous users. Building on these works, [40] show that, through appropriate revenue refunding, a central planner can achieve both the efficiency and equity objectives of sustainable transportation. As in these works, we also consider refunding a proportion of the collected revenues to certain groups of users. However, compared to these works wherein the refund serves as additional money "in the pocket" of users, in our setting, we consider credit-based congestion pricing schemes, wherein the credits provided to users function solely for use of the express lane (e.g., an express lane travel allowance).

CBCP schemes have been explored extensively in the literature with a focus on *tradable* credit schemes wherein additional credits may be purchased [52, 75, 72] or earned by exhibiting desirable travel behavior (e.g., choosing not to travel, travelling at an off-peak period, or travelling by a more sustainable mode of transportation) [73, 55]. Moreover, these credits typically have monetary value beyond paying for access to

priced facilities (e.g., for using public transit) and have the potential to improve the equitable distribution of benefits from congestion pricing [72, 45]. However, the equity improvements from tradable CBCP schemes typically result from lower-income users that reduce travel to sell their allocated credits or use them on driving alternatives [55, 45]. In contrast to these works, we consider non-tradable CBCP schemes in which eligible users receive a fixed budget of credits to access the express lane as an alternative to driving on slower GP lanes. In particular, in the setting studied in this work, credits cannot be traded nor provide value for anything other than for paying express lane tolls, which mirrors the operation of the San Mateo 101 Express Lanes Project. Furthermore, under non-tradable CBCP schemes, eligible users will avail of a fast and reliable mode of travel, i.e., the express lane, which is in contrast to tradable CBCP schemes, wherein low-income users generally use less convenient modes of travel to obtain monetary gains from selling excess credits.

Since travel credits provide no value to users beyond paying for express lane tolls, our work is also closely related to the design of artificial currency mechanisms [36, 42, 31], which have found many applications [61, 9, 25]. However, in contrast to these applications that consider a single-economy, wherein artificial currencies serve as the only medium of transfer to avail resources, we consider a mixed-economy wherein only a fraction of the users receive travel credits (artificial currencies) to use the express lane.

From a methodological viewpoint, as in prior studies that characterize user equilibria with heterogeneous users in congestion games [19, 74] and study equilibrium existence under alternative pricing schemes, such as tradable credits [76, 70], we also investigate the properties of equilibria induced by CBCP schemes. However, we note that CBCP equilibria differ markedly from prior equilibrium notions in the congestion pricing literature (see Section 3.3). Beyond an equilibrium characterization of CBCP schemes, we also develop a bi-level optimization framework to optimize over CBCP schemes to achieve particular societal objectives. While bi-level optimization [16] has been extensively studied in traffic routing contexts, e.g., second-best tolling [68, 47, 50, 57] and revenue maximization [8],our bi-level framework involves optimizing over both tolls and budgets rather than only road tolls.

3 Model

In this section, we introduce the basic definitions of traffic flow for a multi-lane freeway (Section 3.1), the operation of CBCP schemes and corresponding user costs (Section 3.2), and the Nash equilibrium notion we study in this work (Section 3.3).

3.1 Preliminaries

We study the problem of designing CBCP schemes to route heterogeneous users with different VoTs in a multi-lane highway section. The highway consists of one express lane that can be tolled while the remaining general purpose (GP) lanes must remain untolled. Without loss of generality, we model the freeway section as a two-edge Pigou network consisting of a source vertex s, a destination vertex d, and two directed edges $e \in \{1, 2\}$ between the source and destination vertices, where the first edge ($e = 1$) denotes the express lane while the second edge ($e = 2$) corresponds to the GP lanes without tolls. We note that modeling all GP lanes as a single edge in the traffic network is without loss of generality, as we focus on equilibrium formation in this work; hence, these lanes are indistinguishable from the lens of a user as none of these lanes have any tolls. Furthermore, to model the travel times on each edge e in this traffic network, we consider a flow-dependent travel-time (latency) function $l_e : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, which maps x_e, the traffic flow rate on edge e, to the travel time $l_e(x_e)$. As in prior literature on traffic routing, we assume that the function l_e, for both edges e, is differentiable, convex, and monotonically increasing.

Users make trips in the transportation network over T periods (e.g., days) over which the CBCP scheme is run (see Section 3.2) and belong to a finite set of discrete groups characterized by their (i) level of income and (ii) value of time. Let $\mathcal{G}$ denote the set of all user groups, wherein users, based on their income, are subdivided into two categories, *eligible* and *ineligible*, depending on their eligibility to receive travel credit (budget) to use the express lane as pre-determined by a central planner. We let $\mathcal{G}_E$ and $\mathcal{G}_I$ denote the sets of eligible and ineligible user groups, respectively. Furthermore, a user in a group $g \in \mathcal{G}$ at each period $t \in [T]$ has a VoT $v_{t,g}$, which captures users' willingness to pay for travel time savings, i.e., their trade-off between travel time and money. The total travel demand of a user group g is given by d_g, which represents the flow of users in group g to be routed in the network at each period. We assume that the travel demand for each

user group stays fixed across time, as is consistent with weekday rush hour traffic wherein users commute to and from work. For the simplicity of exposition, we normalize d_g to one for all groups g and note that our results naturally extend to the general travel demands setting.

A flow pattern $\mathbf{y} = \{y_{e,t}^g : e \in \{1, 2\}, t \in [T], g \in \mathcal{G}\}$ specifies for each user group g the amount of flow $y_{e,t}^g \geq 0$ routed on edge e at period t. The resulting flows must satisfy the user demand at each period t, i.e.,

$$\sum_{e=1}^{2} y_{e,t}^g = 1, \quad \text{for all } g \in \mathcal{G}, t \in [T].$$

Furthermore, we represent the edge flows corresponding to the flow pattern $\mathbf{y}$ by the vector $\mathbf{x} = \{x_{e,t} : e \in \{1, 2\}, t \in [T]\}$, where

$$x_{e,t} = \sum_{g \in \mathcal{G}} y_{e,t}^g, \quad \text{for all } e \in \{1, 2\}, t \in [T].$$

3.2 Credit-based Congestion Pricing Schemes and User Optimization

A credit-based congestion pricing (CBCP) scheme is characterized by a tuple $(\boldsymbol{\tau}, B)$, where $\boldsymbol{\tau} \in \mathbb{R}_{\geq 0}^T$ represents the vector of tolls on the express lane over T periods, and B is the travel credit (budget) given to eligible users to use the express lane over the T periods. Both eligible and ineligible users pay the corresponding toll when using the express lane; however, ineligible users pay for the tolls out-of-pocket while eligible users pay for the tolls using their available budget. For simplicity, we assume that eligible users never spend out-of-pocket to use the express lane and thus use it only if they have sufficient credits. Such an assumption is consistent with congestion pricing schemes in real-world traffic networks, wherein lower-income users, i.e., those that constitute the eligible group, are less likely to spend money out-of-pocket to use tolled roads. However, we note that this modeling assumption can readily be relaxed to the setting where eligible users can also spend out-of-pocket, and we defer a thorough treatment of this setting to future research. We also note that the finiteness of the time horizon T is crucial to successfully deploying a CBCP scheme to improve express lane access for eligible users as the ratio $\frac{B}{T}$ represents the per-period budget for these users to use the express lane. Finally, we mention that, as in the traffic routing literature [32, 45], we focus on single-occupancy vehicles in this work and defer the consideration of high-occupancy alternatives (e.g., carpools) to future work.

In response to a given CBCP scheme $(\boldsymbol{\tau}, B)$, users are assumed to be selfish and thus choose whether to use the express lane over the T periods to minimize their cumulative travel cost. We now present the individual user optimization problems for both the ineligible and eligible users and note that these differ from each other as we are in a mixed economy setting, wherein ineligible users pay for the express lane tolls out-of-pocket while eligible users use travel credit to use the express lane.

Ineligible Users: Since ineligible users spend out-of-pocket to use the express lane, the cumulative travel cost for ineligible users is assumed to be a linear function of their travel time and tolls, which is a commonly used modeling assumption [15, 28]. In particular, given a CBCP scheme $(\boldsymbol{\tau}, B)$ and a vector of edge flows $\mathbf{x}$, the individual optimization of an ineligible user in a group $g \in \mathcal{G}_I$ is given by

$$\mu^{g*}(\mathbf{x}, \boldsymbol{\tau}, B) = \min_{\mathbf{z}^g \in \mathbb{R}_{\geq 0}^{2 \times T}} \sum_{t \in [T]} \sum_{e \in [2]} (v_{t,g} l_e(x_{e,t}) + \mathbb{1}_{e=1} \tau_t) z_{e,t}^g, \tag{1a}$$

$$\text{s.t.} \quad z_{1,t}^g + z_{2,t}^g = 1, \quad \forall t \in [T] \tag{1b}$$

where (1a) is the travel cost objective of the ineligible users and (1b) are user allocation constraints at each period $t \in [T]$. Observe that Problem (1a)-(1b) is a linear program as the decision variables are given by the vector $\mathbf{z}^g = \{z_{e,t}^g : e \in \{1, 2\}, t \in [T]\}$, which correspond to the actions of an infinitesimal user and thus does not influence the edge flow vector $\mathbf{x}$. Here, we denote the decision variables for any user in group $g \in \mathcal{G}_I$ by the vector $\mathbf{z}^g$ to distinguish it from the cumulative flow of all users in $g \in \mathcal{G}_I$ given by $\mathbf{y}^g = \{y_{e,t}^g : e \in \{1, 2\}, t \in [T]\}$ and note that the decision variables $z_{e,t}^g \in [0, 1]$ can be interpreted as the probability that a user in group g uses edge e at period t (or the fraction of users in group g on edge e at

period t). Further, for succinctness, for any feasible $\mathbf{z}^g$, we denote the travel cost for users in a group $g \in \mathcal{G}_I$ as $\mu_{\mathbf{z}^g}(\mathbf{x}, \boldsymbol{\tau}, B) = \sum_{t=1}^{T} \sum_{e=1}^{2} (v_{t,g} l_e(x_{e,t}) + \mathbb{1}_{e=1} \tau_t) z_{e,t}^g$, the travel cost on edge e at period t for ineligible users in a group $g \in \mathcal{G}_I$ as $\mu_{e,t}^g(\mathbf{x}, \boldsymbol{\tau}, B) = v_{t,g} l_e(x_{e,t}) + \mathbb{1}_{e=1} \tau_t$, and let $\mu^{g*}(\mathbf{x}, \boldsymbol{\tau}, B)$ denote the minimum travel cost.

Optimal Solution of Problem (1a)-(1b): By the separability of the travel cost function and the constraints across periods, the optimal solution of Problem (1a)-(1b) corresponds to travel decisions that are independent across periods. In particular, ineligible users choose the edge with the minimum travel cost at each period, i.e., at each period t ineligible users choose the edge $e \in \arg\min_{e \in \{1,2\}} \{v_{t,g} l_e(x_{e,t}) + \mathbb{1}_{e=1} \tau_t\}$. Observe that the travel behavior of ineligible users at any given period is akin to the well-studied model of heterogeneous users in non-atomic congestion games [15, 28, 74], wherein users choose routes with the minimum travel cost.

Eligible Users: On the other hand, since eligible users only utilize travel credit to use the express lane (and do not spend any money out-of-pocket), the cumulative travel cost for a user belonging to a group $g \in \mathcal{G}_E$ only consists of the travel time component of the cost of the ineligible users. As a result, given a CBCP scheme $(\boldsymbol{\tau}, B)$ and a vector of edge flows $\mathbf{x}$, the individual optimization of an eligible user in a group $g \in \mathcal{G}_E$ is given by

$$\mu^{g*}(\mathbf{x}, \boldsymbol{\tau}, B) = \min_{\mathbf{z}^g \in \mathbb{R}_{\geq 0}^{2 \times T}} \quad \sum_{t \in [T]} \sum_{e \in [2]} v_{t,g} l_e(x_{e,t}) z_{e,t}^g, \tag{2a}$$

$$\text{s.t.} \quad z_{1,t}^g + z_{2,t}^g = 1, \forall t \in [T] \tag{2b}$$

$$\sum_{t \in [T]} z_{e,t}^g \tau_t \leq B, \tag{2c}$$

where (2b) are user allocation constraints at each period $t \in [T]$, and (2c) is the budget constraint, which ensures that no user spends more travel credits than the provided allowance. As with the ineligible users, Problem (2a)-(2c) is a linear program and we denote the travel cost for users in a group $g \in \mathcal{G}_E$ as $\mu_{\mathbf{z}^g}(\mathbf{x}, \boldsymbol{\tau}, B) = \sum_{t=1}^{T} \sum_{e=1}^{2} v_{t,g} l_e(x_{e,t}) z_{e,t}^g$ for any feasible $\mathbf{z}^g = \{z_{e,t}^g : e \in \{1,2\}, t \in [T]\}$ and let $\mu^{g*}(\mathbf{x}, \boldsymbol{\tau}, B)$ denote the minimum travel cost.

Optimal Solution of Problem (2a)-(2c): To characterize the optimal solution of Problem (2a)-(2c), we first note that due to the budget Constraint (2c), the travel decisions of eligible users are coupled across the periods unlike the ineligible users. In particular, the structure of the optimal solution of Problem (2a)-(2c) depends on the notion of a travel *bang-per-buck* ratio, which we define below.

Definition 1 (Travel Bang-Per-Buck Ratio). Consider a CBCP scheme $(\boldsymbol{\tau}, B)$ with a corresponding edge flow vector $\mathbf{x}$. Then, the travel bang-per-buck ratio at any period t for an eligible user in a group $g \in \mathcal{G}_E$ is given by $\frac{v_{t,g}(l_2(x_{2,t}) - l_1(x_{1,t}))}{\tau_t}$.

In other words, the travel bang-per-buck ratio represents the ratio between the travel time savings for an eligible user when using the express lane at period t to the corresponding toll at that period. We now present a characterization of the structure of the optimal solution of Problem (2a)-(2c).

Lemma 1 (Optimal Solution of Individual Optimization Problem of Eligible Users). *Consider a CBCP scheme $(\boldsymbol{\tau}, B)$ and an edge flow $\mathbf{x}$. Then, for any eligible user in a group $g \in \mathcal{G}_E$ the optimal solution to Problem (2a)-(2c) is such that they spend their budget B in using the express lane at different periods in the descending order of the travel bang-per-buck ratios.*

Lemma 1 establishes that eligible users use the express lane at different periods in the descending order of the travel bang-per-buck ratios until their budget is exhausted. This result on the optimal solution to Problem (2a)-(2c) is akin to prior characterizations of the optimal solution to users' individual optimization problems in matching markets [4] as well as Fisher markets with additional knapsack constraints [39]. We refer to Appendix C.1 for a proof of Lemma 1 and note that this result follows directly through an investigation of the first order necessary and sufficient KKT conditions of Problem (2a)-(2c).

3.3 CBCP Equilibria

We evaluate the efficacy of a CBCP scheme in achieving particular societal scale goals of a central planner based on the induced Nash equilibria (see Section 6). To this end, in this section, we present the Nash equilibrium notion, which we term a CBCP equilibrium, studied in this work. In particular, given a CBCP scheme $(\boldsymbol{\tau}, B)$, a flow pattern $\mathbf{y}$ is a CBCP equilibrium if no user can reduce their travel cost through a unilateral deviation, as is formalized by the following definition.

Definition 2 (CBCP Equilibrium). For a given CBCP scheme $(\boldsymbol{\tau}, B)$, the flow $\mathbf{y}$, with corresponding edge flows $\mathbf{x}$, is a CBCP $(\boldsymbol{\tau}, B)$-equilibrium if for each ineligible group $g \in \mathcal{G}_I$ with $y_{e,t}^g > 0$,

$$\mu_{e,t}^g(\mathbf{x}, \boldsymbol{\tau}, B) \leq \mu_{e',t}^g(\mathbf{x}, \boldsymbol{\tau}, B), \quad \text{for all } e' \in \{1, 2\}, t \in [T],$$

and for each eligible group $g \in \mathcal{G}_E$ with $y_{e,t}^g > 0$, it holds that $z_{e,t}^{g*} > 0$ for some optimal solution $\mathbf{z}^*$ to Problem (2a)-(2c), i.e.,

$$\mu_{\mathbf{z}^*}^g(\mathbf{x}, \boldsymbol{\tau}, B) \leq \mu_{\mathbf{z}}^g(\mathbf{x}, \boldsymbol{\tau}, B), \quad \text{for all } \mathbf{z} \geq \mathbf{0} \text{ satisfying Constraints (2b)-(2c).}$$

A few comments about this equilibrium notion are in order. First, recall that the equilibrium condition of the ineligible users in Definition 2 corresponds to the optimal solution of Problem (1a)-(1b). Thus, if all users were ineligible, CBCP equilibria reduce to standard non-atomic Nash equilibria with heterogeneous users [15, 28, 74], as their travel decisions are independent across periods. Next, the equilibrium notion in Definition 2 additionally accounts for the preferences of eligible users whose travel decisions are coupled across the periods through their budget constraints. Thus, CBCP equilibria, which simultaneously capture the preferences of both ineligible and eligible users in a mixed-economy setting, differs from prior works that focus on the single-economy setting wherein all users have quasi-linear costs [15, 28, 74]. For a further discussion on the notion of CBCP equilibria, see Appendix A.

4 Properties of CBCP Equilibria

The study of the properties of Nash equilibria, such as equilibrium existence and the uniqueness of edge flows, in traffic routing has been a focal point of transportation and game theory research, as the efficacy of any introduced policy, e.g., travel demand management strategies such as congestion pricing, is generally evaluated based on the corresponding Nash equilibria. In the context of congestion pricing, the Nash equilibrium characterizations have typically focused on the single-economy settings wherein all users have similar objective functions, i.e., all users have quasi-linear costs. For instance, [74] developed a convex program to establish the existence of Nash equilibria under any set of road tolls and show that the corresponding edge flows are unique for strictly convex travel time functions. In contrast to such classical single-economy settings, in this work, we study a mixed economy setting, wherein different groups of users have different objective functions, depending on whether they are eligible or ineligible. In particular, ineligible users have quasi-linear costs while eligible users solve a budget-constrained cost minimization Problem (2a)-(2c).

As a result, in this section, we initiate our study of CBCP schemes in this mixed-economy setting by studying the properties of CBCP equilibria. In particular, we establish the existence of CBCP equilibria and show that, while the equilibria need not be unique, in general, the corresponding edge flow vector $\mathbf{x}$ is unique for any CBCP scheme (Section 4.1). Then, in Section 4.2, we establish that eligible users' values of time are time-invariant, i.e., their values of time are constant across the T periods, then CBCP equilibria can be computed through the solution of a convex program.

4.1 Existence of CBCP Equilibria and Uniqueness of Edge Flows

In this section, we establish the existence of CBCP equilibria and the uniqueness of the corresponding edge flowsunder any CBCP scheme $(\boldsymbol{\tau}, B)$.

To this end, we first show that CBCP equilibria are guaranteed to exist in the general setting when heterogeneous eligible and ineligible users have time-varying VoTs, i.e., their VoTs can vary across the T periods, as is elucidated through the following theorem.

Theorem 1 (Existence of CBCP Equilibria). *For any CBCP scheme $(\boldsymbol{\tau}, B)$, where the tolls $\boldsymbol{\tau} \geq \mathbf{0}$ and budget $B \geq 0$, there exists a CBCP $(\boldsymbol{\tau}, B)$-equilibrium in the setting with heterogeneous eligible and ineligible users whose VoTs can vary over the T periods.*

Theorem 1 establishes that introducing budgets for the eligible users does not preclude the existence of an equilibrium and augments the literature on investigating equilibrium existence under alternatives to standard congestion pricing [76, 70]. Furthermore, Theorem 1 extends equilibrium existence results in the traffic routing literature focusing on the single-economy setting [74], wherein all users have quasi-linear costs, to the mixed-economy setting studied in this workwith both eligible and ineligible users. To prove Theorem 1, we leverage the following variational inequality characterization of CBCP equilibria.

Lemma 2 (Variational Inequality Characterization of CBCP Equilibria). *Consider a CBCP scheme $(\boldsymbol{\tau}, B)$. Then, a flow $\mathbf{y}^* = (y_{e,t}^{g*})$, with corresponding edge flows $\mathbf{x}^*$, is a CBCP $(\boldsymbol{\tau}, B)$-equilibrium if and only if it solves the following variational inequality problem:*

$$\sum_{t=1}^{T} \sum_{e=1}^{2} \left[\sum_{g \in \mathcal{G}_E} v_{t,g} l_e(x_{e,t}^*)(y_{e,t}^g - y_{e,t}^{g*}) + \sum_{g \in \mathcal{G}_I} (v_{t,g} l_e(x_{e,t}^*) + \mathbb{1}_{e=1} \tau_t)(y_{e,t}^g - y_{e,t}^{g*}) \right] \geq 0, \quad \forall \text{ feasible } \mathbf{y} \in \Omega, \quad (3)$$

where the set Ω is described by $\mathbf{y} \geq \mathbf{0}$, $y_{1,t}^g + y_{2,t}^g = 1$ for all $t \in [T]$ and $g \in \mathcal{G}$, and all eligible users satisfy their budget Constraint (2c).

The proof of Lemma 2 is presented in Appendix C.2 and follows from an analogous argument to that in [54], as the variational Inequality (3) is reminiscent of the finite dimensional variational inequalities used to study heterogeneous user equilibria in classical traffic routing settings [54, 74, 70]. However, in contrast to these approaches that consider a single-economy setting, Equation (3) has two separate terms to capture the differing travel costs for the eligible and ineligible users in a mixed-economy setting.

We now use the variational inequality characterization in Lemma 2 to complete the proof of Theorem 1. To do so, we show that the variational Inequality (3) admits a feasible solution, as is elucidated through the following lemma.

Lemma 3 (Feasibility of Variational Inequality). *There exists a solution $\mathbf{y}^*$ to the variational Inequality (3).*

Proof. First, note that the variational Inequality (3) can be expressed in standard form $F(\mathbf{y}^*)^T (\mathbf{y} - \mathbf{y}^*)$ for all feasible $\mathbf{y}$, where $F = (F_{e,t}^g)_{e \in \{1,2\}, t \in [T], g \in \mathcal{G}}$, $F_{e,t}^g(\mathbf{y}^*) = v_{t,g} t_e(x_{e,t}^*)$ for $g \in \mathcal{G}_E$ and $F_{e,t}^g(\mathbf{y}^*) = v_{t,g} t_e(x_{e,t}^*) + \mathbb{1}_{e=1} \tau_t$ for $g \in \mathcal{G}_I$. Then, following standard variational inequality theory [27, 44], a feasible solution $\mathbf{y}^*$ to Equation (3) exists as the feasible set Ω, defined in Lemma 2, is compact and the travel time functions are continuous. $\square$

Lemmas 2 and 3 jointly imply Theorem 1. While Theorem 1 established the existence of CBCP equilibria, we note that CBCP equilibria need not, in general, be unique. For instance, there are settings when two different CBCP equilibria can be obtained by swapping the actions of two groups of users. Despite this non-uniqueness of CBCP equilibria, we now show that the corresponding aggregate equilibrium edge flows are unique for any CBCP scheme $(\boldsymbol{\tau}, B)$, as is elucidated through the following lemma.

Lemma 4 (Uniqueness of Edge Flows Corresponding to CBCP Equilibria). *For any CBCP scheme $(\boldsymbol{\tau}, B)$, where the tolls $\boldsymbol{\tau} \geq \mathbf{0}$ and budget $B \geq 0$, the aggregate edge flow vector $\mathbf{x}$ corresponding to any CBCP $(\boldsymbol{\tau}, B)$-equilibrium is unique.*

For a proof of Lemma 4, see Appendix C.3. We note that the proof of Lemma 4 relies on the strict convexity of the edge travel time functions, assumed in Section 3.1, which is standard in the traffic routing literature [74] and holds for commonly used travel time functions, such as the BPR function [56]. Furthermore, this assumption on the strict convexity of the edge travel time functions is mild compared to other traffic routing settings that require stronger assumptions on the edge travel time functions, e.g., weighted average monotonicity [54, 70], to establish the uniqueness of edge flows.

4.2 Convex Program to Compute CBCP Equilibria

While Theorem 1 established the existence of CBCP equilibria, determining a feasible solution to the variational Inequality (3) may, in general, be challenging. Several numerical methods to compute solutions to variational inequalities exist with provable convergence guarantees [54]; however, these methods often take many steps to converge to a feasible solution. Given the difficulty of computing solutions to variational inequalities [26], in this section, we present a convex program to compute CBCP equilibria in the setting when eligible users have time-invariant VoTs, i.e., for all eligible user groups $g \in \mathcal{G}_E$, $v_{t,g} = v_{t',g} > 0$ for all $t, t' \in [T]$, while ineligible users can, in general, have time-varying VoTs. We defer the question of computing CBCP equilibria in the setting when eligible users' VoTs vary with time to future research and note that the time-invariance assumption on eligible users' VoTs has important practical significance. In particular, since the express lane is likely to be tolled during morning and evening rush hour periods on weekdays, the VoTs of users commuting to and from work are unlikely to differ much between one period and the next, e.g., between subsequent days. Furthermore, the individual optimization Problem (2a)-(2c) for the eligible users can involve quite sophisticated decision-making as eligible users' travel decisions are coupled across periods. Thus, given the lack of variability in users' VoTs between periods and since users may not have complete information on their values of time over the T periods, eligible users may prefer to minimize their total travel time rather than the more complex Objective (2a).

We first present the convex program to compute CBCP equilibria in the setting when eligible users have VoTs that remain constant over the T periods. In particular, given a CBCP scheme $(\boldsymbol{\tau}, B)$, we consider the following convex optimization problem

$$\min_{\mathbf{y} \in \mathbb{R}_{\geq 0}^{2 \times T \times |\mathcal{G}|}} \quad \sum_{t \in [T]} \left[\sum_{e \in [2]} \int_0^{x_{e,t}} l_e(\omega) d\omega + \sum_{g \in \mathcal{G}_I} \frac{y_{1,t}^g \tau_t}{v_{t,g}} \right], \tag{4a}$$

$$\text{s.t.} \quad y_{1,t}^g + y_{2,t}^g = 1, \quad \forall t \in [T], g \in \mathcal{G} \tag{4b}$$

$$\sum_{t \in [T]} \tau_t y_{1,t}^g \leq B, \quad \forall g \in \mathcal{G}_E, \tag{4c}$$

$$\sum_{g \in \mathcal{G}} y_{e,t}^g = x_{e,t}, \quad \forall e \in E, t \in [T], \tag{4d}$$

where (4b) are allocation constraints, (4c) are eligible user budget constraints, and (4d) are edge flow constraints. For conciseness, we denote the set of feasible flows $\mathbf{y}$ satisfying the constraints of the above convex program as Ω. We reiterate here that even though we consider the setting when eligible users have time-invariant VoTs, the VoTs of the ineligible users can vary with time, as in Objective (4a). Problem (4a)-(4d) is akin to the convex program to compute heterogeneous user equilibria given a vector of road tolls [74]. However, as opposed to the convex program in [74] that considers a single-economy setting, Problem (4a)-(4d), which applies to a mixed-economy setting, only has a toll component in the Objective (4a) for ineligible users and instead has a budget Constraint (4c) for eligible users.

We now present the main result of this section, which establishes that any solution of Problem (4a)-(4d) is a CBCP equilibrium.

Theorem 2 (Convex Program for CBCP Equilibrium Computation)**.** *Consider a CBCP scheme $(\boldsymbol{\tau}, B)$ and the setting when the VoTs of all eligible users do not vary with time, i.e., for all $g \in \mathcal{G}_E$, $v_{t,g} = v_{t',g} > 0$ for all $t, t' \in [T]$. Then, the optimal solution $\mathbf{y}^*$ of the convex Program (4a)-(4d) is a CBCP $(\boldsymbol{\tau}, B)$-equilibrium.*

Proof (Sketch). To prove this claim, we derive the first-order necessary and sufficient KKT conditions of Problem (4a)-(4d) and show that these conditions correspond to equilibrium conditions for all users. Given the close connection between Problem (4a)-(4d) and existing convex programs for computing heterogeneous user equilibria [74], the correspondence between the KKT conditions of Problem (4a)-(4d) and the equilibrium conditions for the ineligible users is immediate. To establish that the KKT conditions of Problem (4a)-(4d) correspond to the equilibrium conditions for the eligible users, we fix a group g and consider two cases depending on the value of the dual variable μ_g of Constraint (4c): (i) $\mu_g = 0$, and (ii) $\mu^g > 0$. In both cases, we use the complimentary slackness conditions and the properties of the optimal solution of the individual optimization Problem (2a)-(2c) of eligible users, derived in Lemma 1, to establish an equivalence between

9

the KKT conditions of Problem (4a)-(4d) and the equilibrium conditions for the eligible users, which proves our claim. □

For a complete proof of Theorem 2, see Appendix C.4. The convex Program (4a)-(4d) provides an efficient method to compute CBCP equilibria in the setting when all eligible users have time-invariant VoTs. In particular, this convex program enables a central planner to evaluate the efficacy of different CBCP schemes on certain societal scale objectives by studying the corresponding equilibria (see Section 6). Finally, we note an immediate consequence of Theorem 2, independently of Lemma 4, that the aggregate edge flow vector $\mathbf{x}$ is unique given any CBCP scheme $(\boldsymbol{\tau}, B)$ when the travel time functions are strictly convex in the setting when all eligible users have values of time that do not vary over the T periods.

5   Comparative Statics Analysis of CBCP Equilibria

In this section, we investigate the properties of CBCP equilibria by performing a comparative statics analysis to characterize the changes in the equilibria induced by CBCP schemes given changes in the tolls set on the express lane or budgets distributed to eligible users. An investigation of the comparative statics of CBCP equilibria provides insights regarding the traffic patterns that are likely to be realized under changes to a CBCP scheme through modifications of the road tolls or the distributed budgets. In particular, such an analysis can help guide a central planner regarding the direction in which the tolls or budgets should be adjusted to achieve desired traffic patterns in the system. Furthermore, as we consider a mixed-economy setting, as opposed to classically studied single-economy settings, a comparative statics analysis helps glean insights into how the introduction of budgets to a certain fraction of users influences traffic patterns.

To this end, we initiate our comparative statics analysis of CBCP equilibria by studying the changes in the aggregate equilibrium express lane flows when the express lane toll is increased or decreased (Sections 5.1 and 5.2). We then investigate how an increase or decrease in the budgets distributed to eligible users influences the aggregate equilibrium express lane flows and the corresponding eligible user travel costs (Section 5.3). We note while several of our obtained comparative statics results align with standard economic intuition, we also obtain some counter-intuitive results, e.g., the violation of a natural substitutes condition (see Section 5.2 for a definition), due to the introduction of travel credits for eligible users.

5.1   Aggregate Edge Flows in Response to Toll Changes

In this section, we study the change in the aggregate equilibrium express lane flow at a given period when the toll on the express lane is increased or decreased. In particular, in alignment with economic intuition, we show that an increase in the express lane toll at period t results in a (weak) reduction in the aggregate equilibrium express lane flow at that period.

Lemma 5 (Monotonicty of Edge Flows with Tolls)**.** *Consider two CBCP schemes $(\boldsymbol{\tau}, B)$ and $(\tilde{\boldsymbol{\tau}}, B)$, where $\tilde{\tau}_t > \tau_t$ and $\tilde{\tau}_{t'} = \tau_{t'}$ for all $t' \neq t$. Then, at equilibrium, the aggregate flows on the express lane at period t satisfies $x_{1,t}(\tilde{\boldsymbol{\tau}}) \leq x_{1,t}(\boldsymbol{\tau})$, where $x_{e,t}(\boldsymbol{\tau})$ denotes the equilibrium aggregate flow on edge e at period t under the CBCP scheme with toll $\boldsymbol{\tau}$.*

For a proof of Lemma 5, see Appendix C.5.[1] Lemma 5 establishes a natural (weakly) monotonically decreasing relationship between the aggregate equilibrium flow on the express lane and the corresponding toll at period t. While this monotonicity relation between edge flows and tolls naturally holds in a single-economy setting, wherein all users pay money out-of-pocket, Lemma 5 extends this monotonicity relation to the mixed-economy setting wherein a certain proportion of the users are given travel credits. We reiterate that the result of Lemma 5 mirrors classical economic theory wherein the demand for a resource is (weakly) reduced with an increase in its price.

---

[1] We note that Lemma 5 can also be extended to the setting when the tolls on the express lane are the same at each period, i.e., $\tau_t = \tau_{t'}$ for all $t, t' \in [T]$. In particular, using arguments similar to those used in the proof of Lemma 5, it can be shown that if the toll at each period is increased from τ to $\tilde{\tau}$, then the aggregate flow on the express lane at each period is (weakly) reduced.

5.2 Violation of Substitutes Condition

While the monotonicity relation established in Lemma 5 aligns with standard intuition from economic theory, we now show that, due to the introduction of travel credits for eligible users, a natural substitutes condition may be violated even in the setting when all eligible users have time-invariant values of time. We note that the substitutes condition is fundamental to the study of classical economic theory, as it serves as a critical condition for the existence of market-clearing prices at which the demand for the given resources equals the capacity for those resources [43, 35].

To establish that the substitutes condition does not hold in the traffic routing setting considered in this work, we first recall the substitutes condition from economic theory, which states that an increase in the price of a particular resource cannot result in reduced demand for other resources. Then, in the context of CBCP schemes, the substitutes condition can be stated as follows.

Definition 3 (Substitutes Condition for CBCP Schemes). Consider two CBCP schemes $(\boldsymbol{\tau}, B)$ and $(\tilde{\boldsymbol{\tau}}, B)$, where $\tilde{\tau}_t > \tau_t$ and $\tilde{\tau}_{t'} = \tau_{t'}$ for all $t' \neq t$. Then, the aggregate equilibrium edge flows $\mathbf{x}(\boldsymbol{\tau})$ satisfies the substitutes condition if $x_{1,t'}(\tilde{\tau}) \geq x_{1,t'}(\boldsymbol{\tau})$ holds for all periods $t' \neq t$, where $x_{e,t}(\boldsymbol{\tau})$ denotes the equilibrium aggregate flow on edge e at period t under the CBCP scheme with toll $\boldsymbol{\tau}$.

In particular, Definition 3 states that if the toll on the express lane is increased at period t with the express lane tolls at all other periods kept fixed, then the express lane aggregate flow at any period $t' \neq t$ weakly increases. We now show using a counterexample that the equilibria induced by CBCP schemes do not satisfy this substitutes condition even in the setting when all eligible users have time-invariant values of time, as is elucidated through the following proposition.

Proposition 1 (Violation of Substitutes Property). *Consider two CBCP schemes $(\boldsymbol{\tau}, B)$ and $(\tilde{\boldsymbol{\tau}}, B)$, where $\tilde{\tau}_t > \tau_t$ and $\tilde{\tau}_{t'} = \tau_{t'}$ for all $t' \neq t$. Then, if the eligible users have time-invariant values of time, there exists an instance such that for some period $t' \neq t$, the equilibrium aggregate flows on the express lane satisfies $x_{1,t'}(\tilde{\tau}) < x_{1,t'}(\boldsymbol{\tau})$. Here, $x_{e,t}(\boldsymbol{\tau})$ denotes the equilibrium aggregate flow on edge e at period t under the CBCP scheme with toll $\boldsymbol{\tau}$.*

Proposition 1 establishes that even in the setting when eligible users have time-invariant values of time, i.e., the condition under which CBCP equilibria can be computed using the convex Program (4a)-(4d), the substitutes condition may not hold. While we defer a proof of Proposition 1 to Appendix C.6, a few comments about the violation of the substitutes condition are in order, which also provide insights into the counterexample used to prove this result. To this end, first note that if all users are ineligible, then the substitutes condition trivially holds as ineligible users' travel decisions are independent across the different periods (see Section 3.2). In other words, an increase in the toll at a given period does not influence the aggregate equilibrium express lane flow at another period if all users are ineligible. As a result, the violation of the substitutes condition occurs due to eligible users whose travel decisions are coupled across periods because of their budget Constraint (2c). In particular, in the counterexample used to prove Proposition 1, an increase in the toll at a given period may result in eligible users spending more of their budget to continue using the express lane at that period, resulting in a decreased aggregate edge flow at another period.

We further note that the violation of the substitutes condition has important implications for a central planner seeking to enforce a particular traffic pattern or desired equilibrium flow in the system. For instance, the central planner may seek to maintain certain travel times on the express lane at all periods. As the substitutes condition is critical for the existence of market-clearing prices, the result of Proposition 1 implies that enforcing a desired equilibrium traffic flow in the system (or maintaining certain travel times on the express lane) may not be possible for a central planner.

5.3 Aggregate Edge Flows and Eligible User Travel Costs in Response to Budget Changes

Having studied the influence of the changes in tolls on the equilibria induced by CBCP schemes, we now investigate how an increase or decrease in the budgets distributed to eligible users influences the aggregate equilibrium express lane flows and the corresponding eligible user travel costs. In particular, we first show that the aggregate equilibrium express lane flow at each period (weakly) increases with an increase in the

budget of eligible users. While this result follows as eligible users can use the express lane for more periods with a higher budget, we then show, contrary to intuition, that there are instances when higher eligible user budgets result in increased travel costs for those users even when their values of time are time-invariant.

We begin by establishing the monotonicity relation between the budgets of eligible users and the corresponding aggregate equilibrium express lane flows, as is elucidated through the following lemma.

Lemma 6 (Monotonicty of Edge Flows with Budgets). *Consider two CBCP schemes $(\boldsymbol{\tau}, B)$ and $(\boldsymbol{\tau}, \tilde{B})$, where $\tilde{B} > B$. Then, at equilibrium, the aggregate flows on the express lane at all periods t satisfies $x_{1,t}(B) \leq x_{1,t}(\tilde{B})$, where $x_{e,t}(B)$ denotes the equilibrium aggregate flow on edge e at period t under the CBCP scheme where eligible users receive a budget B.*

Lemma 6 states that the aggregate equilibrium express lane flow is monotonically non-decreasing with the budgets of the eligible users. This result follows as eligible users can use the express lane for more periods with an increased budget. For a proof of Lemma 6, we refer to Appendix C.7 and note that it follows a similar line of reasoning to that in the proof of Lemma 5.

We further note that Lemma 6 aligns with standard economic intuition that an increase in the budgets of eligible users should result in higher express lane flows as eligible users can use the express lane for more periods with a higher budget. Despite this result, we note that an increase in the budgets of eligible users does not necessarily result in reduced travel costs for those users. In particular, we show that there are instances when increasing the budget of eligible users increases their travel costs even when the values of time of eligible users are time-invariant.

Proposition 2 (Non-Monotonicity of Eligible User Travel Costs with Budget Changes). *Consider two CBCP schemes $(\boldsymbol{\tau}, B)$ and $(\boldsymbol{\tau}, \tilde{B})$, where $\tilde{B} > B$. Then, even in the setting when eligible users have time-invariant values of time, there exists an instance such that the eligible users will incur a higher travel cost at the equilibrium induced by the CBCP scheme $(\boldsymbol{\tau}, \tilde{B})$ with the higher budget as compared to that induced by the CBCP scheme $(\boldsymbol{\tau}, B)$ with the lower budget.*

For the counterexample used to prove Proposition 2, see Appendix C.8. We note that the primary reason for the non-monotonic relationship between the change in the budgets and the travel costs of the eligible users, as established in Proposition 2, is that all eligible users (and not just an individual non-atomic user) receive a higher budget. As a result, an increase in the budget for all eligible users creates "competition" between them to use the express lane, driving up the express lane travel times and the travel costs of eligible users.

6 Credit-Based Congestion Pricing Scheme Design

Thus far, we have investigated the properties of equilibria induced by CBCP schemes and developed a convex program to compute CBCP equilibria in the setting when eligible users' values of time are time-invariant. While an analysis of equilibrium properties aids in developing a foundational understanding of the influence of a policy introduced by a central planner on the resulting traffic pattern, a central planner is typically interested in deploying an optimal policy to achieve particular societal goals. For instance, traffic demand management methods, such as congestion pricing, are often deployed to minimize the total travel time of all users in the system. In the context of CBCP schemes on a multi-lane freeway section, central planners are typically concerned with societal goals beyond system efficiency as they provide budgets to lower-income users to balance the travel costs incurred by ineligible and eligible users. Given the range of possible societal objectives of a central planner, in this section, we present a general framework to design optimal CBCP schemes in the setting when eligible users' values of time are time-invariant. In particular, in Section 6.1, we present a bi-level optimization framework for the design of optimal CBCP schemes, wherein the upper-level problem involves a central planner choosing a CBCP scheme $(\boldsymbol{\tau}, B)$ to induce an equilibrium flow given by the solution of the convex Program (4a)-(4d) (the lower-level problem) that optimizes the societal objective of the central planner. Since solving bi-level programs is, in general, computationally challenging, we then present an algorithmic approach based on dense sampling to compute an approximation to the optimal CBCP scheme (Section 6.2). We further motivate applying a dense sampling approach to solve the bi-level program in Section 6.2 by establishing some continuity relations between the resulting equilibria (and aggregate edge flows) and the corresponding toll and budget parameters that characterize a CBCP scheme.

6.1 Bi-Level Optimization Framework

In this section, we present a bi-level optimization framework for designing optimal CBCP schemes to achieve particular societal objectives of a central planner. We focus on the setting when eligible users have time-invariant VoTs, in which case CBCP equilibria can be computed through the solution of the convex Program (4a)-(4d).

To present the bi-level optimization problem of the central planner, we first introduce some notation. In particular, we model the societal objective of a central planner through a cost function $f : \mathbb{R}^{2 \times T \times |\mathcal{G}|} \to \mathbb{R}$, where $f(\mathbf{y})$ denotes the societal cost associated with the flow $\mathbf{y} \geq \mathbf{0}$ that lies in a feasible set Ω defined by Constraints (4b)-(4d). Furthermore, we denote $\mathcal{F}_U \subseteq \mathbb{R}_{\geq 0}^{T+1}$ as the set of feasible CBCP schemes $(\boldsymbol{\tau}, B)$. Then, the goal of the central planner is to find a feasible CBCP scheme $(\boldsymbol{\tau}^*, B^*) \in \mathcal{F}_U$ such that the resulting equilibria $\mathbf{y}(\boldsymbol{\tau}^*, B^*)$ has the lowest societal cost among all feasible CBCP schemes, i.e., $f(\mathbf{y}(\boldsymbol{\tau}^*, B^*)) \leq f(\mathbf{y}(\boldsymbol{\tau}, B))$ for all feasible CBCP schemes $(\boldsymbol{\tau}, B) \in \mathcal{F}_U$, where $\mathbf{y}(\boldsymbol{\tau}, B)$ is an equilibrium flow given by the solution of Problem (4a)-(4d) under the scheme $(\boldsymbol{\tau}, B)$. In particular, the objective of the central planner can be captured through the following bi-level optimization problem

$$\min_{\substack{\mathbf{y} \in \Omega, \\ (\boldsymbol{\tau}, B) \in \mathcal{F}_U}} \quad f(\mathbf{y}(\boldsymbol{\tau}, B)), \tag{5a}$$

$$\text{s.t.} \quad \mathbf{y}(\boldsymbol{\tau}, B) = \text{ Solution of Problem (4a)-(4d)}, \tag{5b}$$

where Constraint (5b) represents the lower-level problem of computing the equilibrium flow given a CBCP scheme $(\boldsymbol{\tau}, B)$. We now present some examples of the feasibility set $\mathcal{F}_U$ and the societal cost function f, which are relevant in CBCP implementations in practical traffic routing contexts, e.g., the San Mateo 101 express Lanes Project.

Example 1 (Feasibility set $\mathcal{F}_U$ as Interval Constraints). As with second-best tolling [68, 47, 50, 57], wherein the sets of allowable tolls on each edge of the traffic network are often represented by interval constraints, the feasibility set $\mathcal{F}_U$ can also be specified by interval constraints on the tolls and budgets. In particular, the set $\mathcal{F}_U$ is such that the toll on the express lane at each period t satisfies $\tau_t \in [\underline{\tau}, \bar{\tau}]$ for some specified bounds $\underline{\tau}, \bar{\tau} \geq 0$ and the budget $B \in [\underline{B}, \bar{B}]$ for some specified bounds $\underline{B}, \bar{B} \geq 0$.

Example 2 (Feasibility set $\mathcal{F}_U$ with Time-Invariant Tolls and Interval Constraints). In practical traffic routing settings, road tolling schemes are often static, with tolls that do not vary over time. In such a setting, the feasibility set $\mathcal{F}_U$ is a subset of the corresponding feasibility set in Example 1, with the additional restriction that the tolls on the express lane additionally satisfy $\tau_t = \tau_{t'}$ for all periods $t, t' \in [T]$.

Example 3 (Societal Cost Function f as a Weighted Combination of Travel Costs of Eligible and Ineligible Users). When deploying a CBCP scheme, a central planner typically accounts for its social welfare effects on all groups of users. A widely studied social welfare function in redistributive market design involves aggregating the utilities (or costs) of all users and weighting users' costs by a social welfare weight, which corresponds to the relative social importance that the central planner places on the welfare of this user as compared to other users [3, 64]. In particular, each user group g can be associated with a social welfare weight λ_g. Then, the cost function f is given by $f(\mathbf{y}(\boldsymbol{\tau}, B)) = \sum_{g \in \mathcal{G}_E} \lambda_g v_g \sum_{t=1}^{T} \sum_{e=1}^{2} l_e(x_{e,t}) y_{e,t}^g + \sum_{g \in \mathcal{G}_I} \lambda_g \sum_{t=1}^{T} \sum_{e=1}^{2} (v_{t,g} l_e(x_{e,t}) + \mathbb{1}_{e=1} \tau_t) y_{e,t}^g$, where the VoT of users in eligible groups is denoted as v_g as their values of time are assumed to be time-invariant. Note that a higher welfare weight for a given user group implies that the central planner is prioritizing those user groups relative to others.

Example 4 (Societal Cost Function f as Revenue Maximization). A common objective for a central planner is revenue maximization [8], in which case the function f can be represented as the negative of the total tolls collected from the ineligible users when they use the express lane, i.e., $f(\mathbf{y}(\boldsymbol{\tau}, B)) = -\sum_{t=1}^{T} \sum_{g \in \mathcal{G}_I} \tau_t y_{1,t}^g$. Note that no revenue is received from eligible users, as these users only utilize the provided travel credits to use the express lane. Furthermore, there is no loss in revenue in providing travel credits to eligible users, as these are eventually recuperated by the central planner when eligible users use the express lane and expend credit (or once the credits expire after the T periods over which the CBCP scheme is run).

Finally, we note that the bi-level Program (5a)-(5b) can also be readily extended to the setting when a central planner imposes feasibility restrictions on the equilibrium flows $\mathbf{y}$, i.e., the flow $\mathbf{y}$ must belong to a

set $\mathcal{F}_L \subseteq \Omega$. Such restrictions on the equilibrium flows can arise when the central planner seeks to maintain a certain quality of service on the express lane by ensuring that the total number of users on that lane does not exceed a specified threshold. For the simplicity of exposition, we do not focus our attention on such feasibility restrictions on the equilibrium flows, and for the remainder of this work, consider the setting when a central planner's objectives are purely specified through a societal cost function f rather than through constraints on the set of feasible flows $\mathbf{y}$.

6.2 Algorithmic Approach to Solve Bi-level Optimization Problem

In the previous section, we showed that a central planner's objective of determining the optimal CBCP scheme to minimize a societal cost function f can be expressed as a bi-level optimization problem. However, solving bi-level programs is, in general, computationally challenging, and, in particular, even bi-level linear programs are NP-hard to approximate up to any constant factor in general [21]. Furthermore, the societal cost function f may be non-linear as it may depend on the edge travel time functions (see Example 3), which are generally non-linear. Given these difficulties, analytically characterizing the optimal solution of the bi-level program is likely not possible. As a result, in this section, we present a *dense sampling* approach to computing an approximate solution for the bi-level Problem (5a)-(5b) and discuss its computational tractability and practical applicability. We further motivate applying the dense sampling approach to solve the bi-level program by establishing continuity relations between the resulting equilibria (and aggregate edge flows) and the corresponding toll and budget parameters that characterize a CBCP scheme. For the simplicity of exposition, we suppose that the feasible set $\mathcal{F}_U$ is given by interval constraints, as in Example 1, and note that our approach is also applicable for a broader range of feasibility sets.

Dense Sampling Approach: To solve the bi-level Problem (5a)-(5b), we discretize the feasible set $\mathcal{F}_U$ given by interval constraints as a grid with a step size of s in each component (in general, the step size can vary across each component). That is, the express lane toll at any period t lies in the set $\mathcal{A}_s = \{\underline{\tau}, \underline{\tau} + s, \ldots, \bar{\tau}\}$ and the eligible user budget lies in the set $\mathcal{B}_s = \{\underline{B}, \underline{B} + s, \ldots, \bar{B}\}$. Further, we let $\mathcal{C}_s$ be the set of all toll and budget combinations $(\boldsymbol{\tau}, B)$ in this discretized gridwith a step size of s.Then, to compute a good solution to the bi-level Problem (5a)-(5b) with a low societal cost, we evaluate the optimal solution of the convex Program (4a)-(4d) for each CBCP scheme $(\boldsymbol{\tau}, B)$ in the set $\mathcal{C}_s$ and return the CBCP scheme with an equilibrium flow with the least societal costamong all CBCP schemes in the set $\mathcal{C}_s$.That is, we return a CBCP scheme $(\boldsymbol{\tau}_s^*, B_s^*) \in \mathcal{C}_s$ with a corresponding equilibrium flow $\mathbf{y}(\boldsymbol{\tau}_s^*, B_s^*)$, such that $f(\mathbf{y}(\boldsymbol{\tau}_s^*, B_s^*)) \leq f(\mathbf{y}(\boldsymbol{\tau}, B))$ for all $(\boldsymbol{\tau}, B) \in \mathcal{C}_s$ with equilibrium flows $\mathbf{y}(\boldsymbol{\tau}, B)$.

A few comments about using dense sampling to solve the bi-level Problem (5a)-(5b) are in order. First, noting that a CBCP scheme $(\boldsymbol{\tau}, B)$ belongs to a $T+1$ dimensional space, the computational complexity of dense sampling scales with the number of points in the discretized grid given by $|\mathcal{C}_s| = O(\frac{(\bar{\tau}-\underline{\tau})^T(\bar{B}-\underline{B})}{s^{T+1}})$. In particular, our dense sampling approach involves solving the convex Program (4a)-(4d) $|\mathcal{C}_s|$ times. In other words, the computational complexity of the dense sampling approach scales exponentially in the step-size s with the number of periods T. However, in practical traffic routing settings, the number of periods T is typically moderately sized, e.g., $T = 30$ if the CBCP scheme is run for a month, and the tolls imposed on the express lanes tend to be static and thus fixed over a certain period. Given the time-invariance of practically deployed tolling schemes, as in the setting in Example 2, the dense sampling approach can be reduced from one over a $T + 1$ dimensional space to one over two dimensions, as the toll must be kept constant on the express lane across all periods. Thus, in the setting where the tolls on the express lane at all periods must be kept constant, our proposed dense sampling approach provides a computationally tractable method to compute an optimal CBCP scheme $(\boldsymbol{\tau}_s^*, B_s^*) \in \mathcal{C}_s$. Furthermore, while several other methods to solve bi-level programs exist, e.g., KKT reformulations [6], or second-order methods [22], dense sampling serves as a clear and transparent methodology for a central planner to evaluate the set of all possible CBCP schemes in the set $\mathcal{C}_s$ and select the one that performs the best, i.e., achieves the least societal cost f. In particular, in practical traffic routing settings, a central planner may prioritize finding an (approximately) optimal CBCP scheme with the least possible societal cost f even at the expense of larger computational runtimes, which further motivates the practicality of the dense sampling approach.

Continuity Properties: While dense sampling provides a method to evaluate the optimal CBCP scheme in a discretized set $\mathcal{C}_s$, such a scheme may be sub-optimal for the feasibility set $\mathcal{F}_U$. To this end, we now present continuity properties of the equilibrium flows (and the aggregate edge flows) in the toll and budget parameters which highlight that performing dense sampling helps achieve approximately optimal solutions by the derived continuity relations. In particular, the continuity relations motivate the efficacy of a dense sampling approach as the equilibrium flows and the corresponding societal cost are unlikely to change much between two subsequent points in $\mathcal{C}_s$ for small step sizes s. As a result, the optimal scheme in the set $\mathcal{C}_s$ serves as a good approximation to the optimal solution to the bi-level Program (5a)-(5b) for a small enough step-size s.

We now present our continuity result that relates both the equilibrium flows and the aggregate edge flows to the corresponding toll and budget parameters that characterize a CBCP scheme, as elucidated through the following lemma. Here, we let $\mathcal{Y}(\boldsymbol{\tau}, B)$ denote the set of equilibrium flows corresponding to the solution to Problem (4a)-(4d) since the equilibrium flows are, in general, non-unique for any given CBCP scheme $(\boldsymbol{\tau}, B)$.[2]

Lemma 7 (Continuity of Equilibrium Flows)**.** *Suppose that all eligible users have time-invariant values of time and the feasible set $\mathcal{F}_U$ is such that the tolls $\boldsymbol{\tau} > \mathbf{0}$. Further, let $\mathcal{Y}(\boldsymbol{\tau}, B)$ denote the set of solutions to Problem (4a)-(4d) and $\mathbf{x}(\boldsymbol{\tau}, B)$ denote the corresponding unique edge flow. Then, the correspondence $(\boldsymbol{\tau}, B) \implies \mathcal{Y}(\boldsymbol{\tau}, B)$ is upper semi-continuous and locally bounded and $\mathbf{x}(\boldsymbol{\tau}, B)$ is a continuous function in $(\boldsymbol{\tau}, B)$ over any open set of toll and budget parameters in $\mathcal{F}_U$. Furthermore, if the set of equilibrium flows $\mathcal{Y}(\boldsymbol{\tau}, B)$ is singleton, i.e., $\mathcal{Y}(\boldsymbol{\tau}, B) = \{\mathbf{y}(\boldsymbol{\tau}, B)\}$, then the equilibrium flow $\mathbf{y}(\boldsymbol{\tau}, B)$ is continuous in $(\boldsymbol{\tau}, B)$ for any open set of toll and budget parameters in $\mathcal{F}_U$.*

Lemma 7 establishes that the correspondence $\mathcal{Y}(\boldsymbol{\tau}, B)$ of equilibrium flows is upper semi-continuous and locally bounded (we refer to Appendix B for definitions of these terms) and the aggregate edge flow $\mathbf{x}(\boldsymbol{\tau}, B)$ is a continuous function in the toll and budget parameters characterizing a CBCP scheme. We prove Lemma 7 through an application of Berge's theorem of the maximum [46] and present a complete proof of this claim in Appendix C.9. Finally, we note an immediate consequence of Lemma 7, which implies that the CBCP scheme found using dense sampling is a good approximation for the optimal solution to the bi-level Program (4a)-(4d) when the function f depends solely on the aggregate edge flows $\mathbf{x}$, i.e., the sum of the flows of all users.

Corollary 1. *Suppose that all eligible users have time-invariant values of time, the feasible set $\mathcal{F}_U$ is such that the tolls $\boldsymbol{\tau} > \mathbf{0}$, and the societal cost function f depends solely on the aggregate edge flows $\mathbf{x}$, i.e., the bi-level optimization Objective (5a) can be expressed as $f(\mathbf{x}(\tau, B))$, where $\mathbf{x}(\boldsymbol{\tau}, B)$ is the edge flow corresponding to the solution of the convex Program (4a)-(4d). Then, if the function f is continuous in the edge flows $\mathbf{x}$ it holds that f is also continuous in $(\boldsymbol{\tau}, B)$.*

The proof of Corollary 1 is immediate from the continuity of function compositions. Observe that an example of a function f that satisfies the condition of Corollary 1 is the total travel time of all users, which can be expressed as $\sum_{t=1}^{T} \sum_{e=1}^{2} x_{e,t}(\boldsymbol{\tau}, B) l_e(x_{e,t}(\boldsymbol{\tau}, B))$. By establishing the continuity of the societal cost function f in $(\boldsymbol{\tau}, B)$, Corollary 1 implies that small changes in the toll and budget parameters will result in small changes in the societal cost f. As a result, for a small enough step size s, the CBCP scheme output by the dense sampling provides a good approximation to the optimal CBCP scheme.

7 Numerical Experiments

We now investigate the influence of CBCP schemes on traffic patterns and study their optimal design through a real-world application study of the San Mateo 101 Express Lanes Project. Our numerical results not only validate our comparative statics analysis results but also show that the optimal CBCP scheme can vary widely based on the central planner's objective when deploying a CBCP scheme, demonstrating that a principled approach using bi-level optimization proposed in this work is key to realizing the benefits of CBCP

---

[2]Given the potential non-uniqueness of equilibrium flows, in the statement of Lemma 7, we use the double arrow " $\implies$ " to denote a correspondence, which is a map that associates every point in the domain of the correspondence to a subset in its range. Furthermore, we present formal definitions of an upper semi-continuous and locally bounded correspondence mentioned in Lemma 7 in Appendix B.

schemes. In this section, we first present the numerical implementation details for our experiments and our procedure to calibrate the model parameters of the multi-lane freeway segment, such as the edge travel time functions and users' values of time (Section 7.1). Next, in Section 7.2, we present sensitivity results that demonstrate the variation in the travel time and proportion of users on the express lane with the tolls and budgets, reflective of real-world multi-lane freeway segments with express lanes. Then, in Section 7.3, we apply the dense sampling approach in Section 6.2 to solve the bi-level optimization Problem (5a)-(5b) to obtain optimal CBCP schemes under different societal objectives. Finally, we discuss policy implications of this work in Section 7.4.

7.1   Model Calibration and Implementation Details

In this section, we describe the implementation details for our experiments and the method used to calibrate the edge travel time functions, user demand, and the user values of time based on a real-world application study of the San Mateo 101 Express Lanes Project. We note that as part of this project, San Mateo County is constructing 22 miles of express lanes on the US 101 highway in the San Francisco Peninsula, which connects Santa Clara and San Mateo counties with the City and County of San Francisco. For our experiments, we modelled the northbound portion of the US 101 highway involved in this project. We further obtained data regarding the usage of the newly opened express lane relative to the untolled general purpose (GP) lanes in San Mateo.

Travel time calibration;   As in the US 101 highway, which has one express lane and three GP lanes, we set up our two-edge Pigou network model, wherein the first edge represents the express lane and the second edge corresponds to the three GP lanes without tolls (see Section 3.1 for a discussion on this model simplification). To calibrate the travel time functions on the two edges, we queried average hourly weekday travel speed and vehicle flow data from September 2 - September 30, 2019 from Caltrans' Performance Measurement System (PeMS) database [13], depicted in Figure 1. Using this data, we then fit the parameters of the commonly used BPR travel time function [56], depicted in the blue curve in Figure 1, defined as

$$l_e(x_e) = \xi_e \left(1 + a \left(\frac{x_e}{\kappa_e}\right)^b\right),$$

where a, b are constants, ξ_e is the free-flow travel time on edge e, and κ_e is the capacity, i.e., the number of users beyond which the travel time on the edge rapidly increases, of edge e. The calibrated parameter values are as follows: $a = 0.2, b = 6, \xi_e = 19.4$ minutes, and $\kappa_e = 1650$ veh/hr. Given the non-linearity of the BPR function, to tractably solve the convex Program (4a)-(4d) for large problem instances, we further employed the commonly used piecewise affine approximation of the BPR function [65] depicted in the orange line in Figure 1. In particular, the piecewise linear approximation used is given by

$$l_e(x_e) = l_0 + \begin{cases} 0 & \text{if } x_e \leq \lambda\kappa_e \\ \beta_e(x_{e,t} - \lambda\kappa_e) & \text{otherwise} \end{cases}$$

where l_0 is the height of the horizontal line, β_e is the slope of the second line, and $\lambda\kappa_e$ represents the threshold at which the travel time changes in the piecewise linear approximation. We reiterate that since edge one represents the express lane while edge two represents the three general-purpose lanes, the capacities satisfy $\kappa_2 = 3\kappa_1$ and the slope of the second segment of the piecewise linear approximation satisfies $\beta_2 = \frac{\beta_1}{3}$. Finally, we note that after minimizing the mean squared error between the estimated travel times of the piecewise linear approximation and that obtained from PeMS, the final values for the parameters of the piecewise linear approximation to the BPR function are $l_0 = 19.4$ minutes, $\lambda = 0.786$, $\kappa_1 = 1,650$ veh/hr, and $\beta_1 = 0.1256$. For these parameters, the piecewise linear approximation in orange closely approximates the BPR function in blue in Figure 1, which further motivates the use of a piecewise linear approximation in modeling travel times for our experiments.

Demand profile calibration: The total demand for the 4-lane highway was estimated to be around 8,000 vehicles per hour, about the maximum flow observed in the PeMS data. Furthermore, for our experiments, we
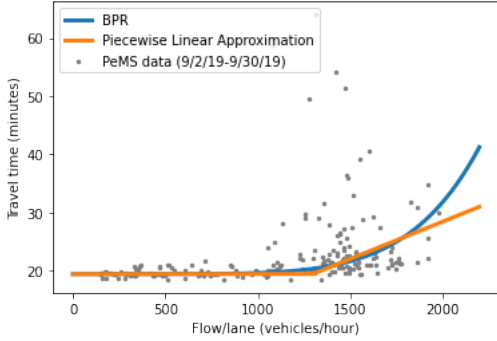
Figure 1: Calibration of a piecewise linear approximation of the travel time function for a single-lane along the San Mateo 101 Express Lanes Project study area.
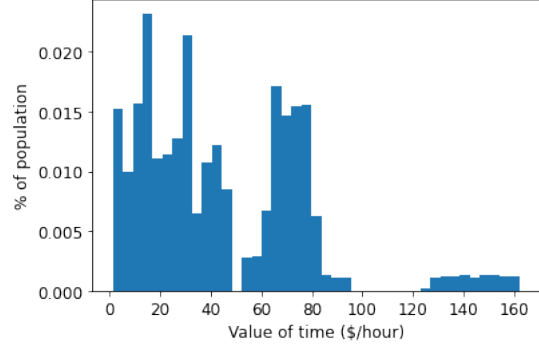

Figure 2: Approximated distribution of the values of time of users (in $/hr) across the San Mateo and Santa Clara Counties.

assumed the CBCP scheme is run over the course of a week with five working days, i.e., $T = 5$. To determine the eligibility of users for free express lane credits, we obtained the distribution of household incomes across users using the 2020 US Census American Community Survey (ACS) annual household earnings estimates for Santa Clara and San Mateo counties [10]. Furthermore, as in the San Mateo 101 Express Lanes Project, we applied a threshold of 200% of the federal poverty limit, resulting in approximately 17% of road users in the eligible group. Next, to approximate the VoT distribution across the users, we used the income distribution from the ACS annual earnings data as a surrogate representation of their VoTs. We reiterate that several other valid representations for users' VoTs are also possible, and we assume proportionality between users' income and their VoT for simplicity. In particular, the VoT of each user at each of the five periods was generated by 1) generating a baseline average VoT for each user using the ACS annual earnings data[3] and 2) for ineligible users, generating deviations from the baseline for each period from a uniform distribution from -0.125 to 0.125 such that the value of time for each user in group g at period t is $v_{t,g} = 1 + v_g \delta_{t,g}$, where v_g is drawn from the probability density function of hourly wages estimated from the ACS data and $\delta_{t,g} \sim \mathcal{U}[-0.125, 0.125]$ for ineligible users and $\delta_{t,g} = 1$ for eligible users. The resulting VoT distribution is displayed in Figure 2, which has a mean of $44/hour and median of $37/hour.

Validation data: In order to validate that the equilibrium edge flows and travel times produced by our model are within the same order of magnitude as in real-world applications, we queried hourly weekday lane-level vehicle flow and travel speed data from September 1 to September 30, 2022 along the first segment of the San Mateo 101 Express Lanes Project that launched in February 2022[4]. Summary statistics are presented in Table 2. On average, during morning peak hours (7-10 am) about 14% of the total flow was on the express lane, resulting in travel time savings of about 33% compared to the GP lanes. Historical data on toll levels for the US 101 Express Lanes Project are not yet available and thus were not included in the validation. However, aggregate toll data was available for the I-680 freeway in Contra Costa County in the East Bay Area [48]. In the first quarter of 2022, the average toll paid on the 20-mile I-680 express lane peaked at about $4.80 per trip for travel time savings of about 2.6 minutes (13%).

---

[3]The ACS annual household income data reports the estimated percentage of the population in each county in each of 10 income groups (less than $10,000, $10,000 to $14,999, ..., $150,000 to $199,999, and $200,000 or more). The probability density function of individual hourly wages, which is used as a proxy for users' VoTs, across the study population was estimated from this data by attributing the middle of each annual income interval to each of the corresponding groups and dividing by total work hours in a year (i.e., 40 hours/week × 52 weeks). Furthermore, since the ACS annual earnings data distinguishes between types of households, including family and non-family households, we divided the estimated wage for family households by two in order to more closely represent *individual* wage levels. The resulting data, consisting of the population size at each hourly wage level, was used to estimate a step-wise probability distribution function for the values of time of users in the case study of the San Mateo 101 Express Lanes Project.

[4]The first of two phases of the San Mateo 101 Express Lanes Project was completed in February, 2022, rolling out the first 6-mile segment of express lane in the Southernmost portion of San Mateo County. Phase two, consisting of the remaining segment of the express lane is scheduled to open in early 2023

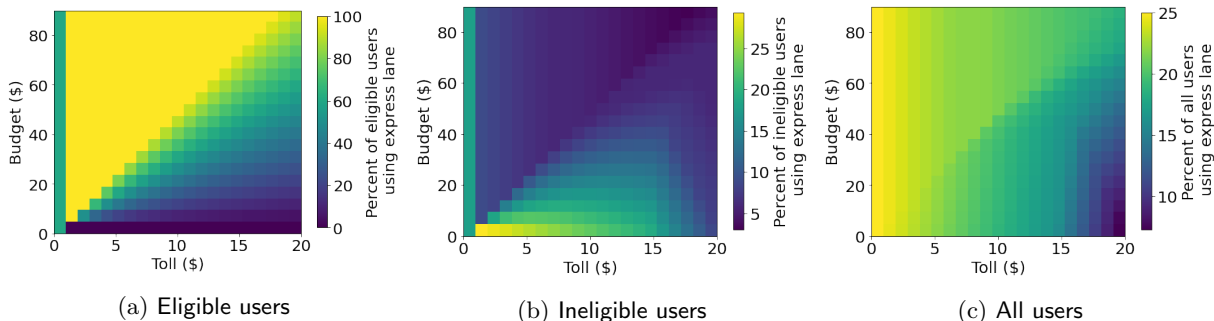(a) Eligible users   (b) Ineligible users   (c) All users

Figure 3: Percentage of users on the express lane corresponding to the optimal solution of the convex Program (4a)-(4d) for different toll and budget combinations. Here the tolls range from \$0 to \$20, with \$1 increments, and budgets range from \$0 to \$90, with \$5 increments.

Implementation details We ran our experiments on an i5-3570 processor with 32 GB of RAM and our corresponding implementation is available at https://anonymous.4open.science/r/CBCP-DB80 We used the Python Gurobi Optimizer (`gurobipy` version 9.5.2) to solve the Convex Program (4a)-(4d) with the above travel time and VoT parameters. On average, it took about 9 seconds to solve this convex program for a given toll and budget combination.

7.2 Express Lane Usage and User Travel Times at CBCP Equilibria

In this section, we present the variation in the travel time and proportion of users on the express lane as the express lane tolls and the budgets distributed to eligible users are varied. In particular, we focus on the setting when eligible users have time-invariant VoTs, and the tolls are the same across the five periods, as in Example 2, over which the CBCP scheme is run. Further, we discretize the tolls to lie between \$0 to \$20, with \$1 increments, and budgets to lie between \$0 to \$90, with \$5 increments, and compute the solution to the convex Program (4a)-(4d) at each of the toll and budget combinations in the discretized grid. The resulting distributions of equilibrium lane choices and travel times corresponding to the optimal solution of Problem (4a)-(4d) are presented in Figures 3 and 4, respectively.

Express Lane Usage: As seen in Figure 3c, users are split evenly across lanes for \$0 tolls, with one-quarter of all users on the express lane and the remaining three-quarters on the three GP lanes. This observation aligns with equilibrium formation in congestion games without tolls, wherein all users traveling between the same origin and destination incur the same travel time. Further, the proportion of eligible users using the express lane ranges from 0% when the budget is \$0 to 100% when the budget exceeds the total cost of tolls over the five periods (i.e., for a toll τ and budget B it holds that $5\tau \leq B$), as reflected by the yellow portion in Figure 3a. On the other hand, the share of ineligible users on the express lane is at a maximum of 29% at the smallest non-zero toll of \$1 and \$0 budget and decreases with either increasing toll or increasing eligible user budget (see Figure 3b). From Figure 3c, we also observe that the overall share of users on the express lane monotonically decreases (increases) with increasing toll (budget) values, reaching a minimum of 7% at the maximum toll of \$20 and \$0 budget. These monotonic relations between the toll and budget parameters and the corresponding proportion of users on the express lane aligns with our comparative statics analysis results in Section 5 (see Lemmas 5 and 6). In addition, the proportion of all users using the express lane smoothly varies with the change in the toll and budget parameters, which aligns with the continuity relation established in Section 6.2 (see Lemma 7).

User Travel Times: From Figures 4a and 4b, we observe that the travel times on the express and GP lanes decrease and increase, respectively, with the overall share of users on the express lane. Further, the travel time savings on the express lane increases monotonically with increasing tolls, with a maximum of about 14.8 minutes (a 43% difference) with a \$20 toll and \$0 budget.

We also note that the overall express lane usage and travel time savings depicted in Figure 3 are comparable to the data obtained from Caltrans' PeMS database [13] for US 101 express lanes in September 2022.

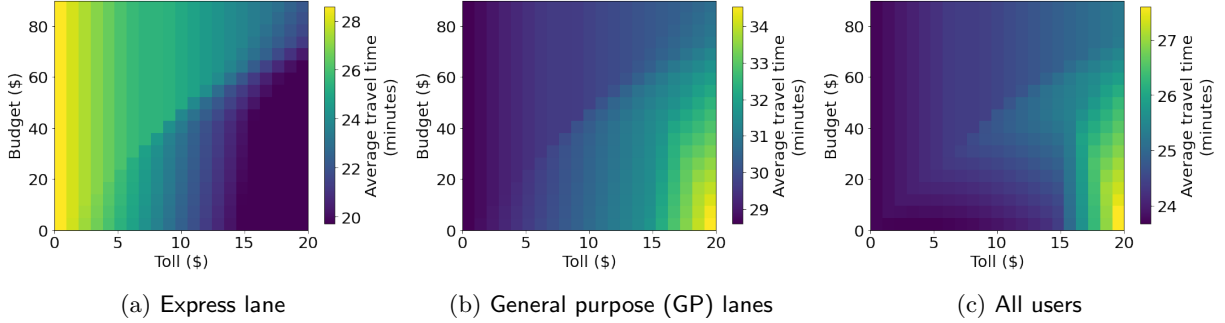| (a) Express lane | (b) General purpose (GP) lanes | (c) All users |

Figure 4: Average travel times corresponding to the optimal the optimal solution of the convex Program (4a)-(4d) for different toll and budget combinations. Here the tolls range from \$0 to \$20, with \$1 increments, and budgets range from \$0 to \$90, with \$5 increments.

In particular, for CBCP schemes with a \$16 toll and budgets between \$0 and \$10, the travel time savings in Figure 3 are about 39%, and express lane usage is around 15%, comparable to that of US 101 express lanes (about 33% and 14%, respectively). Further, the valuation of \$16 for about 12.5 minutes of time savings (about \$1.28/minute) is also comparable to that of the I-680 highway in the San Francisco Bay Area (about \$1.85/minute). Thus, our numerical results reflect real-world multi-lane highways with express lanes.

7.3 Optimal CBCP Schemes

We now study the design of optimal CBCP schemes for the case study of the San Mateo 101 Express Lanes Project for a well-studied societal objective (i.e., the Pareto weighted combination of different cost (or welfare) measures) in the redistributive market design literature [3, 64]. In particular, we consider the societal objective

$$f_{\boldsymbol{\lambda}}(\mathbf{y}(\tau, B)) = \lambda_E \sum_{g \in \mathcal{G}_E} v_g \sum_{t=1}^{T} \sum_{e=1}^{2} l_e(x_{e,t}) y_{e,t}^g - \lambda_R \sum_{g \in \mathcal{G}_I} \sum_{t=1}^{T} \tau_t y_{1,t}^g$$
$$+ \lambda_I \sum_{g \in \mathcal{G}_I} \sum_{t=1}^{T} \sum_{e=1}^{2} (v_{t,g} l_e(x_{e,t}) + \mathbb{1}_{e=1} \tau_t) y_{e,t}^g,$$

Table 1 presents the optimal CBCP schemes that induce an equilibrium with the minimum societal cost for each Pareto weight $\boldsymbol{\lambda}$ and lists the proportion of users on the express lane and corresponding average travel times under the optimal scheme. Furthermore, Figure 5 displays the distribution of the objective function (see Section 7.3) for each of nine different Pareto weighted schemes in Table 1. From Table 1, we observe that the optimal CBCP scheme can vary widely based on the central planner's objective, thus demonstrating that a principled approach using bi-level optimization proposed in this work is key to realizing the benefits of CBCP schemes. For instance, if the central planner solely optimizes for the travel costs of the eligible users, i.e., $\boldsymbol{\lambda} = (1, 0, 0)$, then the optimal CBCP scheme involves providing high budgets (for eligible users to use the express lane) and setting high tolls (to push most ineligible users out of using the express lane). On the other hand, when optimizing for toll revenues, the optimal CBCP scheme corresponds to providing no budgets and setting a slightly lower toll of \$15 (to incentivize enough eligible users to use the express lane).

We also observe from Table 1 that as the Pareto weight corresponding to eligible users is increased from 1 to 10 while keeping $\lambda_I, \lambda_R = 1$, the optimal budget remains \$0 while the optimal toll decreases from \$13 to \$10. A \$0 budget is optimal for Pareto weights $\lambda_E \leq 10$ (with $\lambda_I, \lambda_R = 1$), as eligible users have much lower VoTs due to their lower incomes (see Appendix 7.1) and thus their travel costs do not have a high enough weight relative to that of ineligible users. However, any increase in λ_E beyond 10 results in optimal CBCP schemes with increasing budgets, as eligible users' travel costs are given a higher weight (or priority). Further, note from Table 1 that increasing λ_E from 10 to 12 increases the proportion of eligible users on the

Table 1: Optimal CBCP schemes for various Pareto weights with the corresponding travel times (TTs) and fraction of users on the express lane

Weights $(\lambda_E, \lambda_I, \lambda_R)$	Optimal CBCP τ	B	% using express lane Overall	Eligible	Ineligible	Average TT Express	GP
(1, 0, 0)	19	90	19	95	3	22.1	30.3
(0, 1, 0)	0	0	25	60	18	28.2	28.2
(0, 0, 1)	15	0	16	0	19	19.4	31.4
(1, 1, 1)	13	0	17	0	21	20.3	30.9
(5, 1, 1)	11	0	18	0	22	21.5	30.5
(10, 1, 1)	10	0	19	0	23	22.1	30.3
(11, 1, 1)	10	15	19	30	17	22.3	30.2
(12, 1, 1)	11	45	19	82	7	22.6	30.1
(15, 1, 1)	13	55	19	85	6	22.2	30.3

express lane from 0% to 82%, corresponding to a 21% decrease in travel costs for these users (see Figure 6d in Appendix D.2).

7.4 Policy Implications

Our results have several implications for designing optimal CBCP schemes in practice, e.g., for the San Mateo 101 Express Lanes Project. First, since the optimal CBCP scheme can vary widely based on the central planner's objective (see Table 1), our results show that a principled approach using bi-level optimization is key to realizing the benefits of CBCP. Further, our experimental results suggest that CBCP schemes, at the expense of reduced toll revenues, can help alleviate the equity concerns of congestion pricing and this potential benefit to eligible users outweighs the negative impacts to ineligible users. In particular, compared to the optimal CBCP scheme for $\boldsymbol{\lambda} = (1, 1, 1)$, the optimal CBCP scheme for $\boldsymbol{\lambda} = (11, 1, 1)$ results in an equilibrium flow at which the toll revenues decrease by about 37% while the average travel cost for eligible and ineligible users decreases by 10% and about 1%, respectively (see Figure 5 in Appendix D.2). The reduction in the average travel costs of the ineligible users follows due to the reduced travel time on the GP lanes, on which most ineligible users travel. Further, the optimal CBCP scheme corresponding to $\boldsymbol{\lambda} = (11, 1, 1)$ amounts to a budget of $15 every five days, which is a much higher credit allocation than the one-time $100 allocation currently provided in the San Mateo Community Benefits Program. Hence, our results recommend a more frequent and recurring allocation of credits, e.g., every month, as this credit budget is likely to be used up in a matter of weeks or months, after which the equity benefits of CBCP will be foregone.

While the lowest-income users (i.e., those eligible to receive credits) stand to gain from CBCP, we reiterate, as noted in section 7.2, that the worst-off users under a CBCP scheme are the ineligible users with the lowest VoTs. In particular, these users incur higher travel costs due to the imposition of tolls that increases the travel time on GP lanes, yet typically do not have high enough VoTs to merit paying for the express lanes. Given that the lowest VoT users without credits will bear the brunt of the travel time increases on GP lanes, we believe that a multi-level budget structure, in which credits are allocated to users as some function of their income (and/or other eligibility criteria), may result in a more equitable distribution of the benefits of CBCP.

Beyond the aforementioned CBCP policy implications and recommendations of our work, the success of a CBCP scheme in achieving improved equity outcomes significantly hinges on the consideration of high-occupancy modes of travel, which is beyond the scope of this work. In particular, many express lanes, including those on the US 101 in San Mateo, are actually HOT lanes, in which HOVs (i.e., carpools of two, three, or more persons, shuttles, buses) are also permitted to access the express lane.With free or discounted access for HOVs, tolls may also reduce the total flow in the network by incentivizing a shift from single occupant vehicles (SOVs) to HOVs. Depending on the prevailing rate of carpooling and willingness to adopt carpooling, the equilibrium tolls and budgets are likely to differ from those in the present study.

Finally, as part of the Community Transportation Benefits Program of the San Mateo 101 Express Lanes Project, eligible residents also have the choice of receiving $100 in public transit fares instead of express lane

(a) $(\lambda_E, \lambda_I, \lambda_T) = (1, 0, 0)$

(b) $(\lambda_E, \lambda_I, \lambda_T) = (0, 1, 0)$

(c) $(\lambda_E, \lambda_I, \lambda_T) = (0, 0, 1)$

(d) $(\lambda_E, \lambda_I, \lambda_T) = (1, 1, 1)$

(e) $(\lambda_E, \lambda_I, \lambda_T) = (5, 1, 1)$

(f) $(\lambda_E, \lambda_I, \lambda_T) = (10, 1, 1)$

(g) $(\lambda_E, \lambda_I, \lambda_T) = (11, 1, 1)$

(h) $(\lambda_E, \lambda_I, \lambda_T) = (12, 1, 1)$

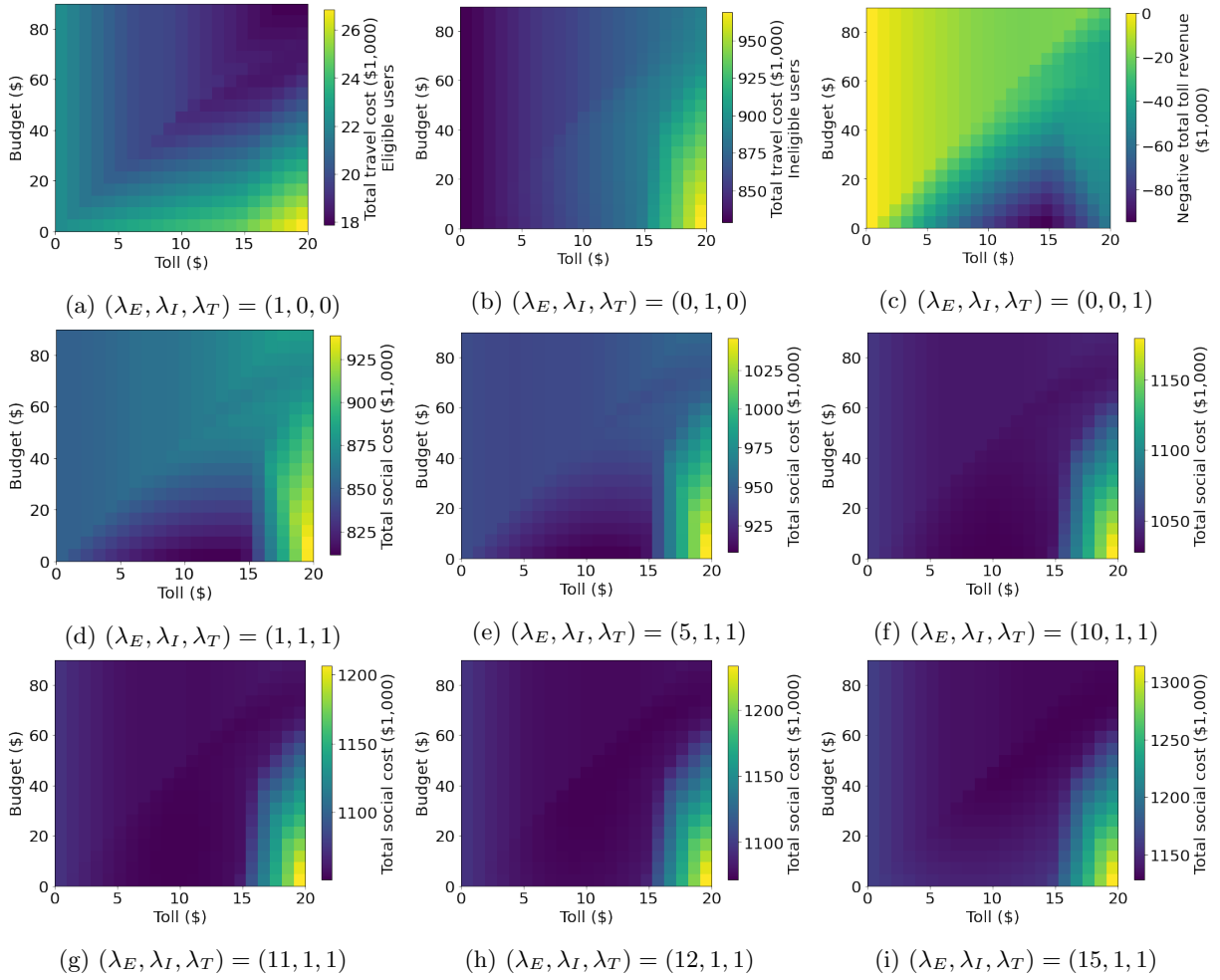(i) $(\lambda_E, \lambda_I, \lambda_T) = (15, 1, 1)$

Figure 5: Objective function values (see Section 7.3) with varying pareto weights on the total travel costs of eligible and ineligible users and the negative total toll revenue

credits. Such an option benefits low-income residents who may not be able to drive or are willing to forego driving (e.g., by using another mode), which may further reduce vehicle flows and amplify the benefits of the CBCP scheme. We defer the thorough examination of CBCP schemes in the context of HOVs and more complex credit schemes to future research.

8 Conclusion and Future Work

In this paper, we studied CBCP schemes, analogous to those implemented in the San Mateo 101 Express Lanes Project, to route heterogeneous users with different VoTs in a multi-lane highway. Since CBCP schemes involve giving travel credits to a fraction of the users, we introduced a new mixed economy model, wherein eligible users receive travel credits while the remaining ineligible users spend money out-of-pocket to use the express lane. In this mixed economy setting, we first investigated the effect of CBCP schemes on the resulting traffic patterns by characterizing the properties of the corresponding Nash equilibria. In particular, we established the existence of CBCP equilibria and, in the setting when eligible users have time-invariant values of time, presented a convex program to compute CBCP equilibria. We further performed a comparative statics analysis to establish monotonicity relations between the equilibria induced by CBCP schemes in response to a change in the tolls set on an express lane or budgets distributed to eligible users. Next, we introduced a bi-level optimization framework to design optimal CBCP schemes to achieve particular societal objectives and presented a dense sampling approach to compute an approximation to the optimal CBCP scheme. Finally, we presented numerical experiments based on an application study of the San Mateo 101 Express Lanes Project, which demonstrated that a principled approach using bi-level optimization proposed in this work is key to realizing the benefits of CBCP schemes.

There are several directions for future research. First, it would be worthwhile to investigate whether equilibria can be computed efficiently in the more general setting when eligible users' VoTs are time-varying. Next, it would be interesting to explore methods to improve the computational complexity of the dense sampling approach. For instance, as opposed to the general dense sampling approach developed in this paper, which is independent of the societal cost function f, using properties of f combined with the comparative statics analysis results (see Section 5) can be used to narrow the search space, i.e., only a subset of points $\mathcal{C}_s$ need to be searched over. Furthermore, several model extensions, e.g., considering time-varying travel demand and the sensitivity of departure time choices to tolling schedules, are of interest to further mirror the real-world operation of express lanes. Moreover, including HOVs and incorporating mode choices would further improve understanding of the role of modal shift incentives in optimal CBCP. Lastly, investigating more general budget allocation structures, e.g., as a step-wise function of incomes, is a promising direction for further improving the equity of CBCP.

References

[1] Jeffrey L. Adler and Mecit Cetin. 2001. A direct redistribution model of congestion pricing. *Transportation Research Part B: Methodological* 35, 5 (2001), 447–460.

[2] Mohammad Akbarpour, Eric Budish, Piotr Dworczak, and Scott Duke Kominers. 2022. An Economic Framework for Vaccine Prioritization. In *Proceedings of the 23rd ACM Conference on Economics and Computation* (Boulder, CO, USA) *(EC '22)*. Association for Computing Machinery, New York, NY, USA, 1181. https://doi.org/10.1145/3490486.3538241

[3] Mohammad Akbarpour, Piotr Dworczak, and Scott Duke Kominers. 2020. *Redistributive allocation mechanisms*. GRAPE Working Papers 40. GRAPE Group for Research in Applied Economics. https://ideas.repec.org/p/fme/wpaper/40.html

[4] Saeed Alaei, Pooya Jalaly Khalilabadi, and Eva Tardos. 2017. Computing Equilibrium in Matching Markets. In *Proceedings of the 2017 ACM Conference on Economics and Computation* (Cambridge, Massachusetts, USA) *(EC '17)*. Association for Computing Machinery, New York, NY, USA, 245–261. https://doi.org/10.1145/3033274.3085150

[5] Richard Arnott, André de Palma, and Robin Lindsey. 1994. The Welfare Effects of Congestion Tolls with Heterogeneous Commuters. *Journal of Transport Economics and Policy* 28, 2 (1994), 139–161.

[6] Navid Azizan Ruhi, Krishnamurthy Dvijotham, Niangjun Chen, and Adam Wierman. 2018. Opportunities for price manipulation by aggregators in electricity markets. *IEEE Transactions on Smart Grid* 9, 6 (2018), 5687–5698.

[7] Dimitris Bertsimas, Vivek F. Farias, and Nikolaos Trichakis. 2011. The Price of Fairness. *Operations Research* 59, 1 (2011), 17–31. https://doi.org/10.1287/opre.1100.0865 arXiv:https://doi.org/10.1287/opre.1100.0865

[8] Luce Brotcorne, Martine Labbé, Patrice Marcotte, and Gilles Savard. 2001. A Bilevel Model for Toll Optimization on a Multicommodity Transportation Network. *Transportation Science* 35, 4 (2001), 345–358. https://doi.org/10.1287/trsc.35.4.345.10433 arXiv:https://doi.org/10.1287/trsc.35.4.345.10433

[9] Eric Budish. 2011. The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes. *Journal of Political Economy* 119, 6 (2011), 1061–1103. https://doi.org/10.1086/664613

[10] US Census Bureau. 2021. 2021 American Community Survey (ACS) 1-Year Estimates - S2001: Earnings in the Past 12 Months (In 2021 Inflation-Adjusted Dollars). https://data.census.gov/cedsci/table?q=S2001&g=0500000US06081,06085&tid=ACSST1Y2021.S2001

[11] Ricardo Cano. 2021. S.F. is considering downtown 'congestion pricing.' Here's how much it would cost. SF Chronicle. Retrieved October 14, 2022 from https://www.sfchronicle.com/sf/article/S-F-is-considering-downtown-congestion-16336449.php

[12] Frederic Charlier Carlos F. Pardo. 2022. Bogotá Tries to Make Congestion Pricing Flexible and Equitable. Streets Blog USA. Retrieved October 21, 2022 from https://usa.streetsblog.org/2022/10/17/bogota-tries-to-make-congestion-pricing-flexible-and-equitable/

[13] Chao Chen and Pravin Varaiya. 2002. *Freeway Performance Measurement System (Pems)*. Ph.D. Dissertation. AAI3082138.

[14] Steve Chien and Alistair Sinclair. 2011. Convergence to approximate Nash equilibria in congestion games. *Games and Economic Behavior* 71, 2 (2011), 315–327. https://doi.org/10.1016/j.geb.2009.05.004

[15] Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. 2003. Pricing Network Edges for Heterogeneous Selfish Users. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing* (San Diego, CA, USA) *(STOC '03)*. Association for Computing Machinery, New York, NY, USA, 521–530. https://doi.org/10.1145/780542.780618

[16] Benoît Colson, Patrice Marcotte, and Gilles Savard. 2007. An overview of bilevel optimization. *Annals of operations research* 153, 1 (2007), 235–256.

[17] Daniele Condorelli. 2013. Market and non-market mechanisms for the optimal allocation of scarce resources. *Games and Economic Behavior* 82 (2013), 582–591.

[18] Metropolitan Transportation Corporation. 2022. Bay Area Express Lanes. https://mtc.ca.gov/operations/traveler-services/bay-area-express-lanes

[19] Stella C. Dafermos. 1972. The Traffic Assignment Problem for Multiclass-User Transportation Networks. *Transportation Science* 6, 1 (1972), 73–87. https://doi.org/10.1287/trsc.6.1.73 arXiv:https://doi.org/10.1287/trsc.6.1.73

[20] Carlos F. Daganzo. 1995. A pareto optimum congestion reduction scheme. *Transportation Research Part B: Methodological* 29, 2 (1995), 139–154.

[21] Stephan Dempe, Vyacheslav Kalashnikov, Gerardo A. Prez-Valds, and Nataliya Kalashnykova. 2015. *Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks*. Springer Publishing Company, Incorporated.

[22] Robert Dyro, Edward Schmerling, Nikos Arechiga, and Marco Pavone. 2022. Second-Order Sensitivity Analysis for Bilevel Optimization. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 9166–9181.

[23] Editor. 2022. San Mateo 101 Express Lanes Open with a First-of-its-Kind Equity Program. Everything South City. Retrieved April 27, 2022 from https://everythingsouthcity.com/2022/04/san-mateo-101-express-lanes-open-with-a-first-of-its-kind-equity-program/

[24] Jonas Eliasson and Lars-Göran Mattsson. 2006. Equity effects of congestion pricing: Quantitative methodology and a case study for Stockholm. *Transportation Research Part A: Policy and Practice* 40, 7 (2006), 602–620. https://doi.org/10.1016/j.tra.2005.11.002

[25] Ezzat Elokda, Carlo Cenedese, Kenan Zhang, John Lygeros, and Florian Dörfler. 2022. CARMA: Fair and efficient bottleneck congestion management with karma. https://doi.org/10.48550/ARXIV.2208.07113

[26] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. 2004. The complexity of pure Nash equilibria. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*. 604–612.

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

[28] Lisa Fleischer, Kamal Jain, and Mohammad Mahdian. 2004. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In *IEEE Symposium on Foundations of Computer Science*. https://doi.org/10.1109/FOCS.2004.69

[29] Kurtulus Gemici, Elias Koutsoupias, Barnabé Monnot, Christos H. Papadimitriou, and Georgios Piliouras. 2019. Wealth Inequality and the Price of Anarchy. In *International Symposium on Theoretical Aspects of Computer Science (Leibniz International Proceedings in Informatics, Vol. 126)*. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 31:1–31:16.

[30] Jose A Gómez-Ibáñez and Kenneth A Small. 1994. *Road pricing for congestion management: A survey of international practice*. Vol. 210. Transportation Research Board.

[31] Artur Gorokh, Siddhartha Banerjee, and Krishnamurthy Iyer. 2021. From Monetary to Nonmonetary Mechanism Design via Artificial Currencies. *Mathematics of Operations Research* 46, 3 (2021), 835–855. https://doi.org/10.1287/moor.2020.1098

[32] Xiaolei Guo and Hai Yang. 2010. Pareto-improving congestion pricing and revenue refunding with multiple user classes. *Transportation Research Part B: Methodological* 44, 8 (2010), 972–982.

[33] Jonathan D. Hall. 2018. Pareto improvements from Lexus Lanes: The effects of pricing a portion of the lanes on congested highways. *Journal of Public Economics* 158 (2018), 113–125. https://doi.org/10.1016/j.jpubeco.2018.01.003

[34] James Hannan. 2016. *4. APPROXIMATION TO RAYES RISK IN REPEATED PLAY*. Princeton University Press, Princeton, 97–140. https://doi.org/doi:10.1515/9781400882151-006

[35] John W. Hatfield and Paul R Milgrom. 2005. Matching with contracts. *American Economic Review* 95, 4 (2005), 913–935.

[36] Aanund Hylland and Richard Zeckhauser. 1979. The Efficient Allocation of Individuals to Positions. *Journal of Political Economy* 87, 2 (1979), 293–314. https://doi.org/10.1086/260757

[37] S. Jaensirisak, M. Wardman, and A. D. May. 2005. Explaining Variations in Public Acceptability of Road Pricing Schemes. *Journal of Transport Economics and Policy* 39, 2 (2005), 127–153.

[38] O. Jahn, R. Möhring, A. Schulz, and N. Stier-Moses. 2005. System-Optimal Routing of Traffic Flows with User Constraints in Networks with Congestion. *Operations Research* 53, 4 (2005), 600–616. https://doi.org/10.1287/opre.1040.0197

[39] Devansh Jalota, Marco Pavone, Qi Qi, and Yinyu Ye. 2021. Fisher Markets with Linear Constraints: Equilibrium Properties and Efficient Distributed Algorithms. arXiv:2106.10412 [cs.GT]

[40] Devansh Jalota, Kiril Solovey, Karthik Gopalakrishnan, Stephen Zoepf, Hamsa Balakrishnan, and Marco Pavone. 2021. When Efficiency Meets Equity in Congestion Pricing and Revenue Refunding Schemes. In *Equity and Access in Algorithms, Mechanisms, and Optimization*. Association for Computing Machinery, New York, NY, USA, Article 9, 11 pages. https://doi.org/10.1145/3465416.3483296

[41] Devansh Jalota, Kiril Solovey, Matthew Tsao, Stephen Zoepf, and Marco Pavone. 2022. Balancing Fairness and Efficiency in Traffic Routing via Interpolated Traffic Assignment. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems* (Virtual Event, New Zealand) *(AAMAS '22)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 678–686.

[42] Ian A Kash, Eric J Friedman, and Joseph Y Halpern. 2007. Optimizing scrip systems: Efficiency, crashes, hoarders, and altruists. In *Proceedings of the 8th ACM conference on Electronic commerce*. 305–315.

[43] Alexander S. Kelso and Vincent P. Crawford. 1982. Job Matching, Coalition Formation, and Gross Substitutes. *Econometrica* 50, 6 (1982), 1483–1504. http://www.jstor.org/stable/1913392

[44] David Kinderlehrer and Guido Stampacchia. 2000. *An Introduction to Variational Inequalities and Their Applications*. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898719451 arXiv:https://epubs.siam.org/doi/pdf/10.1137/1.9780898719451

[45] Kara M. Kockelman and Sukumar Kalmanje. 2005. Credit-based congestion pricing: a policy proposal and the public's response. *Transportation Research Part A: Policy and Practice* 39, 7 (2005), 671–690. https://doi.org/10.1016/j.tra.2005.02.014

[46] David M. Kreps. 2012. *Microeconomic Foundations I: Choice and Competitive Markets*. Number 9890 in Economics Books. Princeton University Press.

[47] Martine Labbé, Patrice Marcotte, and Gilles Savard. 1998. A Bilevel Model of Taxation and Its Application to Optimal Highway Pricing. *Management Science* 44, 12-part-1 (1998), 1608–1622. https://doi.org/10.1287/mnsc.44.12.1608 arXiv:https://doi.org/10.1287/mnsc.44.12.1608

[48] Bay Area Express Lanes. 2022. *I-680 Contra Costa Express Lanes Performance Report - 1st Quarter 2022: January-March*. Technical Report. Metropolitan Transportation Comission. https://mtc.ca.gov/sites/default/files/documents/2022-08/Q1-2022-BAIFA-Express-Lanes-Quarterly-Performance-Report.pdf

[49] Tore Langmyhr. 1999. Understanding innovation: the case of road pricing. *Transport Reviews* 19, 3 (1999), 255–271.

[50] Torbjörn Larsson and Michael Patriksson. 1998. *Side Constrained Traffic Equilibrium Models—Traffic Management Through Link Tolls*. Springer US, Boston, MA, 125–151. https://doi.org/10.1007/978-1-4615-5757-9_7

[51] David Levinson. 2010. Equity Effects of Road Pricing: A Review. *Transport Reviews* 30, 1 (2010), 33–57. https://doi.org/10.1080/01441640903189304 arXiv:https://doi.org/10.1080/01441640903189304

[52] Weijia (Vivian) Li, Kara M. Kockelman, and Yantao Huang. 2021. Traffic and Welfare Impacts of Credit-Based Congestion Pricing Applications: An Austin Case Study. *Transportation Research Record* 2675, 1 (2021), 10–24. https://doi.org/10.1177/0361198120960139 arXiv:https://doi.org/10.1177/0361198120960139

[53] T Litman. 2003. London Congestion Pricing: Implications for Other Cities. Victoria Transport Policy Institute, Victoria, BC, Canada.

[54] A. Nagurney. 2000. A multiclass, multicriteria traffic network equilibrium model. *Mathematical and Computer Modelling* 32, 3 (2000), 393–411. https://doi.org/10.1016/S0895-7177(00)00142-4

[55] Yu Marco Nie. 2015. A New Tradable Credit Scheme for the Morning Commute Problem. *Networks and Spatial Economics* 15 (2015), 719–741. https://doi.org/10.1007/s11067-013-9192-8

[56] Bureau of Public Roads. 1964. *Traffic assignment manual*. Technical report, U.S. Dept. of Commerce, Urban Planning Division. https://books.google.com/books?id=t9XhugEACAAJ

[57] Michael Patriksson and R. Tyrrell Rockafellar. 2002. A Mathematical Model and Descent Algorithm for Bilevel Traffic Management. *Transportation Science* 36, 3 (2002), 271–291. https://doi.org/10.1287/trsc.36.3.271.7826 arXiv:https://doi.org/10.1287/trsc.36.3.271.7826

[58] Tyler Patterson and David M Levinson. 2008. Lexus lanes or corolla lanes? Spatial use and equity patterns on the I-394 MnPASS lanes. (2008).

[59] Azi Paybarah. 2019. Congestion Pricing: Mass Transit Savior or Tax on the Working Class? New York Times.

[60] C. Kenneth Poole, Robert W. Jr. Orski. 2000. HOT Lanes: A Better Way to Attack Urban Highway Congestion Transportation. *Regulation* 23 (2000), 15.

[61] Canice Prendergast. 2017. How Food Banks Use Markets to Feed the Poor. *Journal of Economic Perspectives* 31, 4 (November 2017), 145–62. https://doi.org/10.1257/jep.31.4.145

[62] Tim Roughgarden. 2005. *Selfish Routing and the Price of Anarchy*. The MIT Press.

[63] Tim Roughgarden and Éva Tardos. 2002. How Bad is Selfish Routing? *J. ACM* 49, 2 (March 2002), 236–259. https://doi.org/10.1145/506147.506153

[64] Emmanuel Saez and Stefanie Stantcheva. 2016. Generalized Social Marginal Welfare Weights for Optimal Tax Theory. *American Economic Review* 106, 1 (January 2016), 24–45. https://doi.org/10.1257/aer.20141362

[65] M. Salazar, M. Tsao, I. Aguiar, M. Schiffer, and M. Pavone. 2019. A Congestion-aware Routing Scheme for Autonomous Mobility-on-Demand Systems. In *European Control Conference*. Naples, Italy. /wp-content/papercite-data/pdf/Salazar.Tsao.ea.ECC19.pdf

[66] Y. Sheffi. 1985. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall.

[67] Kenneth A Small. 1992. Using the revenues from congestion pricing. *Transportation* 19, 4 (1992), 359–381.

[68] Erik T. Verhoef. 2002. Second-best congestion pricing in general static transportation networks with elastic demands. *Regional Science and Urban Economics* 32, 3 (2002), 281–310. https://doi.org/10.1016/S0166-0462(00)00064-8

[69] William S Vickrey. 1969. Congestion theory and transport investment. *The American Economic Review* 59, 2 (1969), 251–260.

[70] Xiaolei Wang, Hai Yang, Daoli Zhu, and Changmin Li. 2012. Tradable travel credits for congestion management with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review* 48, 2 (2012), 426–437. https://doi.org/10.1016/j.tre.2011.10.007

[71] Martin L. Weitzman. 1977. Is the Price System or Rationing More Efficient in Getting a Commodity to Those Who Need it Most? *Bell Journal of Economics* 8, 2 (1977), 517–524.

[72] Di Wu, Yafeng Yin, Siriphong Lawphongpanich, and Hai Yang. 2012. Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks. *Transportation Research Part B: Methodological* 46, 9 (2012), 1273–1287.

[73] Ling-Ling Xiao, Hai-Jun Huang, and Ronghui Liu. 2015. Tradable credit scheme for rush hour travel choice with heterogeneous commuters. *Advances in Mechanical Engineering* 7, 10 (2015), 1687814015612430. https://doi.org/10.1177/1687814015612430 arXiv:https://doi.org/10.1177/1687814015612430

[74] Hai Yang and Hai-Jun Huang. 2004. The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. *Transportation Research Part B: Methodological* 38, 1 (2004), 1–15. https://doi.org/10.1016/S0191-2615(02)00074-7

[75] Hai Yang and Xiaolei Wang. 2011. Managing network mobility with tradable credits. *Transportation Research Part B: Methodological* 45, 3 (2011), 580–594. https://doi.org/10.1016/j.trb.2010.10.002

[76] Hai Yang and Xiaolei Wang. 2011. Managing network mobility with tradable credits. *Transportation Research Part B: Methodological* 45, 3 (2011), 580–594. https://EconPapers.repec.org/RePEc:eee:transb:v:45:y:2011:i:3:p:580-594

A   Further Discussion of CBCP Equilibrium Notion

The notion of CBCP equilibria, introduced in Definition 2 captures the travel decisions of both eligible and ineligible users. However, we do remark that, as with prior equilibrium notions in non-atomic congestion games [15, 28, 74], CBCP equilibria requires certain informational assumptions to be realized. For instance, all users require information on the resulting edge flows to solve their respective individual optimization problems, i.e., Problem (1a)-(1b) for the ineligible users and Problem (2a)-(2c) for the eligible users. Furthermore, due to the coupling of the eligible users' travel decisions across periods, eligible users additionally require information on their values of time over the T periods to make informed travel decisions. As a result, from the perspective of the eligible users, the Nash equilibrium notion in Definition 2 can be interpreted as an equilibrium concept wherein eligible users can look back in hindsight and, based on the realized edge flows $\mathbf{x}$ and their values of time over the T periods, determine if they could have benefited through a unilateral deviation by using the express lane in a different set of periods. We reiterate that no such coupling of travel decisions across periods holds for ineligible users.

Despite the need for certain informational assumptions, we note that the equilibrium notion introduced in Definition 2 provides a reasonable approximation into how both groups of users would make travel decisions as it captures the objectives and preferences of both eligible and ineligible users. Thus, as with prior general equilibrium studies, we focus our attention on fundamentally characterizing properties of the Nash equilibrium concept in Definition 2 to understand the influence of CBCP schemes on traffic outcomes in practice. The question of whether CBCP equilibria can be realized is beyond the scope of this work, and we refer to [14, 34] for a discussion of methods used to study whether the players of a game reach an equilibrium.

Finally, we also note that without loss of generality it suffices to focus on CBCP equilibrium flows $\mathbf{y}$ such that for eligible (ineligible) user groups $g \in \mathcal{G}_E$ ($g \in \mathcal{G}_I$), $y_{e,t}^g = z_{e,t}^{g*}$ for some optimal solution $\mathbf{z}^*$ to Problem (2a)-(2c) (Problem (1a)-(1b)), as all users in a given group incur the same travel cost at any CBCP equilibrium.

B   Definitions

In this section, we present the definitions of locally bounded, upper semi-continuous, and lower semi-continuous correspondences.

Definition 4 (Locally Bounded Correspondence). A correspondence $\phi : X \implies Y$ is locally bounded if for ever $x \in X$, there exists an $\epsilon(x) > 0$ and a bounded set $Y(x) \subseteq Y$ such that $\phi(x') \subseteq Y(x)$ for all x' that are less than an ϵ distance from x.

Definition 5 (Upper Semi-Continuous Correspondence). A correspondence $\phi : X \implies Y$ is upper semi-continuous if, whenever $\{x^n\}$ is a sequence in X with limit x and $\{y^n\}$ is a sequence in Y such that $y^n \in \phi(x^n)$ for all n and $\lim_n y^n$ exists, then this limit point is an element of $\phi(x)$.

Definition 6 (Lower Semi-Continuous Correspondence). A correspondence $\phi : X \implies Y$ is lower semi-continuous if for every $x \in X$, the sequence $\{x^n\}$ in X with limit x, and $y \in \phi(x)$, we can find for all sufficiently large n, i.e., all $n > N$ for some large N, $y^n \in \phi(x^n)$ such that $\lim_n y^n = y$.

C   Proofs

C.1   Proof of Lemma 1

To derive the optimal solution of Problem (2a)-(2c), we first formulate its Lagrangian, which can be written as follows

$$
\mathcal{L} = \sum_{t=1}^{T}\sum_{e=1}^{2} v_{t,g} l_e(x_{e,t}) z_{e,t}^g + \mu \left(\sum_{t=1}^{T} \tau_t z_{1,t}^g - B \right) - \sum_{t=1}^{T} \lambda_t (z_{1,t}^g + z_{2,t}^g - 1) - \sum_{t=1}^{T}\sum_{e=1}^{2} s_{e,t} z_{e,t}^g,
$$

where μ is the dual variable corresponding to the budget Constraint (2c), λ_t is the dual variable corresponding to the allocation Constraint (2b), and $s_{e,t}$ is the dual variable corresponding to the non-negativity constraint.

We now derive the first order necessary and sufficient optimality conditions of Problem (2a)-(2c) by evaluating the derivative of the above defined Lagrangian. For the first edge we obtain that

$$
\frac{\partial \mathcal{L}}{\partial z_{1,t}^g} = v_{t,g} l_1(x_{1,t}) + \mu \tau_t - \lambda_t - s_{1,t} = 0.
$$

From the sign constraint that $s_{1,t} \geq 0$ and the complimentary slackness conditions, the above equation implies that

$$
v_{t,g} l_1(x_{1,t}) + \mu \tau_t \geq \lambda_t, \text{ if } z_{1,t}^g \geq 0,
$$
$$
v_{t,g} l_1(x_{1,t}) + \mu \tau_t = \lambda_t, \text{ if } z_{1,t}^g > 0.
$$

Similarly, we have for the eligible users for the second edge that

$$
v_{t,g} l_2(x_{2,t}) \geq \lambda_t, \text{ if } z_{2,t}^g \geq 0,
$$
$$
v_{t,g} l_2(x_{2,t}) = \lambda_t, \text{ if } z_{2,t}^g > 0.
$$

From the above relations, we observe that $\lambda_t = \min\{v_{t,g} l_1(x_{1,t}) + \mu \tau_t, v_{t,g} l_2(x_{2,t})\}$ for all periods $t \in [T]$. In particular, eligible users in group g use the express lane at periods $t \in [T]$ when $v_{t,g} l_1(x_{1,t}) + \mu \tau_t \leq v_{t,g} l_2(x_{2,t})$. Next, observe that since the dual variable μ is independent of t, this inequality can be rearranged to obtain that $\mu \leq \frac{v_{t,g}(l_2(x_{2,t}) - l_1(x_{1,t}))}{\tau_t}$ for all periods t when eligible users use the express lane and $\mu > \frac{v_{t,g}(l_2(x_{2,t}) - l_1(x_{1,t}))}{\tau_t}$ for all periods t when eligible users do not use the express lane. In other words, users use the express lane at periods t in the descending order of the travel *bang-per-buck* ratio $\frac{v_{t,g}(l_2(x_{2,t}) - l_1(x_{1,t}))}{\tau_t}$ until their budget is exhausted, which proves our claim.

C.2   Proof of Lemma 2

We first prove the forward direction of this claim. Suppose $\mathbf{y}^*$ is a CBCP $(\boldsymbol{\tau}, B)$-equilibrium flow, then we claim that it satisfies the variational inequality problem. To this end, note that at any CBCP $(\boldsymbol{\tau}, B)$-equilibrium flow it must hold for all ineligible users in group $g \in \mathcal{G}_I$ at each period $t \in [T]$ that

$$
\sum_{e=1}^{2} (v_{t,g} l_e(x_{e,t}^*) + \mathbb{1}_{e=1} \tau_t) y_{e,t}^g \geq \sum_{e=1}^{2} (v_{t,g} l_e(x_{e,t}^*) + \mathbb{1}_{e=1} \tau_t) y_{e,t}^{g*}
$$

for all feasible $\mathbf{y}^g = (y^g_{e,t})_{t \in [T], e \in \{1,2\}}$ for user group $g \in \mathcal{G}_I$, since ineligible users will choose the edge that minimizes their travel cost at each period. Similarly, at any Nash equilibrium flow $\mathbf{y}^*$, it holds for all eligible users in group $g \in \mathcal{G}_E$ that

$$\sum_{t=1}^{T} \sum_{e=1}^{2} v_{t,g} l_e(x^*_{e,t}) y^g_{e,t} \geq \sum_{t=1}^{T} \sum_{e=1}^{2} v_{t,g} l_e(x^*_{e,t}) y^{g*}_{et},$$

for all feasible $\mathbf{y}^g = (y^g_{e,t})_{t \in [T], e \in \{1,2\}}$ for user group $g \in \mathcal{G}_E$. Then, summing the above two inequalities for all ineligible and eligible users we obtain that

$$\sum_{g \in \mathcal{G}_E} \sum_{t=1}^{T} \sum_{e=1}^{2} v_{t,g} l_e(x^*_{e,t}) y^g_{e,t} + \sum_{g \in \mathcal{G}_I} \sum_{t=1}^{T} \sum_{e=1}^{2} (v_{t,g} l_e(x^*_{e,t}) + \mathbb{1}_{e=1} \tau_t) y^g_{e,t} \geq$$
$$\sum_{g \in \mathcal{G}_E} \sum_{t=1}^{T} \sum_{e=1}^{2} v_{t,g} l_e(x^*_{e,t}) y^{g*}_{et} + \sum_{g \in \mathcal{G}_I} \sum_{t=1}^{T} \sum_{e=1}^{2} (v_{t,g} l_e(x^*_{e,t}) + \mathbb{1}_{e=1} \tau_t) y^{g*}_{e,t},$$

which can be rearranged to obtain the variational inequality in the statement of the lemma. This proves that any CBCP $(\boldsymbol{\tau}, B)$-equilibrium flow must satisfy the given variational inequality.

To prove the other direction, fix the strategies of all but one eligible user group $g' \in \mathcal{G}_E$ and consider a feasible flow $\mathbf{y}$ such that $\mathbf{y}^g = \mathbf{y}^{g*}$ for all $g \neq g' \in \mathcal{G}_E$, where $\mathbf{y}^g = (y^g_{e,t})_{t \in [T], e \in \{1,2\}}$. Then, it holds by the variational inequality that

$$\sum_{t=1}^{T} \sum_{e=1}^{2} v_{t,g'} l_e(x^*_{e,t}) (y^{g'}_{e,t} - y^{g'*}_{e,t}) \geq 0,$$

which is the equilibrium condition for any user in the eligible group $g' \in \mathcal{G}_E$. Next, fix an ineligible user group $g' \in \mathcal{G}_I$ and some period $t' \in [T]$, and consider a flow $\mathbf{y}$ such that $\mathbf{y}^g = \mathbf{y}^{g*}$ for all $g \neq g' \in \mathcal{G}_I$ and $\mathbf{y}^{g'}_t = \mathbf{y}^{g'*}_t$ for all $t \neq t'$, where $\mathbf{y}^g_t = (y^g_{e,t})_{e \in \{1,2\}}$. Then, it holds by the variational inequality that

$$\sum_{e=1}^{2} (v_{t',g} l_e(x^*_{e,t'}) + \mathbb{1}_{e=1} \tau_{t'}) (y^{g'}_{e,t'} - y^{g'*}_{e,t'}) \geq 0,$$

which is the equilibrium condition for any user in the ineligible group $g' \in \mathcal{G}_I$ at any period $t' \in [T]$, as ineligible users minimize their total travel costs at each period. Since both eligible and ineligible users satisfy their corresponding equilibrium conditions, we have established that any flow $\mathbf{y}^*$ satisfying the variational inequality is a CBCP $(\boldsymbol{\tau}, B)$-equilibrium, which proves our claim.

C.3 Proof of Lemma 4

To prove this claim, we proceed by contradiction. In particular, suppose that there are two equilibrium flows corresponding to a CBCP scheme $(\boldsymbol{\tau}, B)$ with two distinct aggregate edge flows $\mathbf{x}$ and $\tilde{\mathbf{x}}$. Since the aggregate edge flows are distinct, we suppose without loss of generality that $x_{1,t} > \tilde{x}_{1,t}$ for some period t. We now show that the number of ineligible users on the express lane at period t, denoted as $x^{\mathcal{I}}_{1,t}$ and $\tilde{x}^{\mathcal{I}}_{1,t}$, respectively, must satisfy $x^{\mathcal{I}}_{1,t} \leq \tilde{x}^{\mathcal{I}}_{1,t}$, and then use this fact to derive our desired contradiction.

To establish that $x^{\mathcal{I}}_{1,t} \leq \tilde{x}^{\mathcal{I}}_{1,t}$, first note by the equilibrium condition for the ineligible users that a user in group $g \in \mathcal{G}_I$ takes the express lane at period t only when

$$v_{t,g} l_1(x_{1,t}) + \tau_t \leq v_{t,g} l_2(x_{2,t}).$$

Then, noting that $x_{1,t} > \tilde{x}_{1,t}$ and correspondingly that $x_{2,t} < \tilde{x}_{2,t}$ as the user demand is fixed, consider an ineligible user in group $g \in \mathcal{G}_I$ that uses the express lane at equilibrium at period t corresponding to the

edge flows $\mathbf{x}$. It then holds that

$$v_{t,g}l_1(\tilde{x}_{1,t}) + \tau_t \overset{(a)}{<} v_{t,g}l_1(x_{1,t}) + \tau_t,$$
$$\overset{(b)}{\leq} v_{t,g}l_2(x_{2,t}),$$
$$\overset{(c)}{<} v_{t,g}l_2(\tilde{x}_{2,t}),$$

where (a) follows by the assumption that $x_{1,t} > \tilde{x}_{1,t}$, (b) holds as the ineligible user group uses the express lane at period t under the edge flows $\mathbf{x}$, and (c) follows as $x_{2,t} < \tilde{x}_{2,t}$. The above (strict) inequalities imply that if an ineligible user takes the express lane at period t under the edge flow $\mathbf{x}$, then this user also takes the express lane under the edge flow $\tilde{\mathbf{x}}$. That is, we have shown that $x_{1,t}^{\mathcal{I}} \leq \tilde{x}_{1,t}^{\mathcal{I}}$.

Next, let $x_{1,t}^{\mathcal{E}}$ and $\tilde{x}_{1,t}^{\mathcal{E}}$ denote the equilibrium aggregate flow on edge e of eligible users. Then, the above derived relation for ineligible users, together with the fact that $x_{1,t} > \tilde{x}_{1,t}$, implies that $x_{1,t}^{\mathcal{E}} > \tilde{x}_{1,t}^{\mathcal{E}}$. We now show that this relation for the eligible users cannot hold true to derive the desired contradiction.

To see this, recall that under a set of tolls $\boldsymbol{\tau}$ eligible users use the express lane at the periods when the travel bang-per-buck ratio $\frac{v_{t,g}(l_2(x_{2,t}(\boldsymbol{\tau})) - l_1(x_{1,t}(\boldsymbol{\tau})))}{\tau_t}$ is the highest. Then, since the flow $x_{1,t} > \tilde{x}_{1,t}$, the ratio $\frac{v_{t,g}(l_2(x_{2,t}) - l_1(x_{1,t}))}{\tau_t} < \frac{v_{t,g}(l_2(\tilde{x}_{2,t}) - l_1(\tilde{x}_{1,t}))}{\tau_t}$. Further, the fact that $x_{1,t} > \tilde{x}_{1,t}$ implies that there are users from some group g such that $y_{1,t}^g > \tilde{y}_{1,t}^g$, where $\mathbf{y}, \tilde{\mathbf{y}}$ represent the equilibrium flows corresponding to the edge flows $\mathbf{x}, \tilde{\mathbf{x}}$, respectively. Then, by the budget constraint of users in group g it must hold that there is some period t' at which the flow of users in group g on the express lane is lower under the edge flow $\mathbf{x}$, i.e., $y_{1,t'}^g < \tilde{y}_{1,t'}^g$. However, for $\mathbf{y}$ to be an equilibrium, it must hold that for all periods t' such that the flow of users in group g on the express lane is lower than that under $\tilde{\mathbf{y}}$ that the aggregate express lane flow must be higher under $\mathbf{y}$. Note if the aggregate express lane flow at one of the periods t' is lower under $\mathbf{y}$, then users in group g have a profitable deviation at flow $\mathbf{y}$, as they would achieve a lower travel cost when using the express lane at that period as compared to increasing their flow on the express lane at period t (given the change in their travel bang-per-buck ratios).

Next, since the flow of users in group g is lower under the flow $\mathbf{y}$ (as compared to $\tilde{\mathbf{y}}$) at a subset of periods t' when the aggregate express lane flow is higher under $\mathbf{y}$, there must be other groups g' such that their flow is higher at the periods t'. By the above argument for group g, we have that the flow $\mathbf{y}$ can only be an equilibrium if for all groups g'' their express lane flow is higher than under $\tilde{\mathbf{y}}$ at one of the periods with a higher aggregate express lane flow under $\mathbf{y}$. However, there are a subset of periods for which the aggregate express lane flow is higher under $\mathbf{y}$ (which holds by our assumption that $x_{1,t} > \tilde{x}_{1,t}$ at period t) and another set of periods for which the total aggregate express lane flow is lower under $\mathbf{y}$ (which holds by eligible users' budget constraints). Thus, it must be that there is some user group such that their flow is lower under $\mathbf{y}$ at a period when the aggregate express lane flow is lower under $\mathbf{y}$ and is higher at another period when the aggregate express lane flow is higher under $\mathbf{y}$ as compared to $\tilde{\mathbf{y}}$. However, then the flow $\mathbf{y}$ cannot be an equilibrium, which gives us our desired contradiction, proving our claim.

C.4 Proof of Theorem 2

To establish the existence of CBCP $(\boldsymbol{\tau}, B)$-equilibria, we first note that there exists a feasible solution to Problem (4a)-(4d). Next, we can verify that the KKT conditions of Problem (4a)-(4d) precisely correspond to the equilibrium conditions for all users, which establishes the existence of a CBCP $(\boldsymbol{\tau}, B)$-equilibrium.

To prove this claim, we derive the necessary optimality conditions of Problem (4a)-(4d) and show that these conditions imply the equilibrium conditions for both the eligible and ineligible users.

To derive the optimality conditions, we first formulate the following Lagrangian of Problem (4a)-(4d):

$$\mathcal{L} = \sum_{t=1}^{T}\left[\sum_{e=1}^{2}\int_0^{x_{e,t}} l_e(\omega)d\omega + \sum_{g\in\mathcal{G}_I}\frac{y_{1,t}^g \tau_t}{v_{t,g}}\right] + \sum_{g\in\mathcal{G}_E}\mu_g\left(\sum_{t=1}^{T}\tau_t y_{1,t}^g - B\right)$$
$$- \sum_{g\in\mathcal{G}}\sum_{t\in[T]}\lambda_t^g(y_{1,t}^g + y_{2,t}^g - 1) - \sum_{g\in\mathcal{G}}\sum_{t\in[T]}\sum_{e=1}^{2}s_{e,t}^g y_{e,t}^g,$$

where $\sum_{g \in \mathcal{G}} y_{e,t}^g = x_{e,t}$, for all $e \in E, t \in [T]$ as in Constraint (4d). We now derive the first order optimality conditions for both the eligible and ineligible user groups.

Ineligible Users: For an ineligible user group $g \in \mathcal{G}_I$, we consider the following first order optimality condition by evaluating the derivative of the above defined Lagrangian. In particular, for the first edge we obtain that

$$\frac{\partial \mathcal{L}}{\partial y_{1,t}^g} = l_1(x_{1,t}) + \frac{\tau_t}{v_{t,g}} - \lambda_t^g - s_{1,t}^g = 0.$$

From the sign constraint that $s_{1,t}^g \geq 0$ and the complimentary slackness conditions, the above equation implies that

$$l_1(x_{1,t}) + \frac{\tau_t}{v_{t,g}} \geq \lambda_t^g, \text{ if } y_{1,t}^g \geq 0,$$
$$l_1(x_{1,t}) + \frac{\tau_t}{v_{t,g}} = \lambda_t^g, \text{ if } y_{1,t}^g > 0.$$

Similarly, we have for the ineligible users for the second edge that

$$l_2(x_{2,t}) \geq \lambda_t^g, \text{ if } y_{2,t}^g \geq 0,$$
$$l_2(x_{2,t}) = \lambda_t^g, \text{ if } y_{2,t}^g > 0.$$

In other words, the travel cost for an ineligible user at each period t is given by $v_{t,g}\lambda_t^g = \min\{v_{t,g}l_1(x_{1,t}) + \tau_t, v_{t,g}l_2(x_{2,t})\}$, i.e., ineligible users choose the edge corresponding to the minimum travel cost at each period t. As a result, the resulting outcome corresponds to the equilibrium conditions for the ineligible users.

Eligible Users: We now derive the first order necessary optimality conditions of the eligible users, $g \in \mathcal{G}_E$ by evaluating the derivative of the above defined Lagrangian. For the first edge we obtain that

$$\frac{\partial \mathcal{L}}{\partial y_{1,t}^g} = l_1(x_{1,t}) + \mu_g \tau_t - \lambda_t^g - s_{1,t}^g = 0.$$

From the sign constraint that $s_{1,t}^g \geq 0$ and the complimentary slackness conditions, the above equation implies that

$$l_1(x_{1,t}) + \mu_g \tau_t \geq \lambda_t^g, \text{ if } y_{1,t}^g \geq 0,$$
$$l_1(x_{1,t}) + \mu_g \tau_t = \lambda_t^g, \text{ if } y_{1,t}^g > 0.$$

Similarly, we have for the eligible users for the second edge that

$$l_2(x_{2,t}) \geq \lambda_t^g, \text{ if } y_{2,t}^g \geq 0,$$
$$l_2(x_{2,t}) = \lambda_t^g, \text{ if } y_{2,t}^g > 0.$$

Using the above defined relations, the sum of the travel costs for the eligible users over the T periods under the optimal solution to Problem (4a)-(4d) is given by

$$\sum_{t=1}^{T} \lambda_t^g - \mu_g B = \sum_{t=1}^{T} \min\{l_1(x_{1,t}) + \mu_g \tau_t, l_2(x_{2,t})\} - \mu_g B.$$

In the case when $\mu_g = 0$, we have that the travel cost for eligible users in group g satisfies

$$\sum_{t=1}^{T} \min\{l_1(x_{1,t}), l_2(x_{2,t})\} \leq \sum_{t=1}^{T} \sum_{e \in E} l_e(x_{e,t}) \tilde{y}_{e,t}^g$$

31

for any feasible $\tilde{\mathbf{y}}^g$ for $g \in \mathcal{G}_E$ as $\tilde{y}_{1,t}^g + \tilde{y}_{2,t}^g = 1$ for all periods $t \in [T]$. As a result, in the case when $\mu_g = 0$, the optimal solution to Problem (4a)-(4d) corresponds to the equilibrium conditions for the eligible users.

Next, consider the case when $\mu_g > 0$. In this case, by the complimentary slackness conditions it holds that $\sum_{t=1}^{T} \tau_t y_{1,t}^g = B$. Next, suppose without loss of generality that the edge flows are ordered such that $x_{1,1} \leq x_{1,2} \leq \ldots \leq x_{1,T}$. Then, since μ_g is independent of t it must hold that there is some T_1 such that $\mu_g \tau_t + l_1(x_{1,t}) \geq l_2(x_{2,t})$ for $t > T_1$ and that $\mu_g \tau_t + l_1(x_{1,t}) \leq l_2(x_{2,t})$ for $t \leq T_1$ by the monotonicity of the $x_{1,t}$'s and the fact that the total demand $x_{1,t} + x_{2,t}$ is a fixed quantity across the T periods. Here T_1 is such that $\sum_{t=1}^{T_1} \tau_t y_{1,t}^g = B$. That is, a user in group g is routed on edge one at those periods when the ratio of the difference in the travel time to toll ratio on that edge is the lowest, which aligns with the characterization of the optimal solution of eligible users' individual optimization problem in Lemma 1 for the setting when eligible users' values of time are time-invariant. As a result, it follows that the total cost of the eligible users is given by

$$\sum_{t=1}^{T_1} l_1(x_{1,t}) + \sum_{t=T_1}^{T} l_2(x_{2,t}) \leq \sum_{t=1}^{T} \sum_{e=1}^{2} l_e(x_{e,t}) \tilde{y}_{e,t}^g,$$

which holds for any feasible $\tilde{\mathbf{y}}^g$ for $g \in \mathcal{G}_E$. Thus, the optimal solution of Problem (4a)-(4d) corresponds to the equilibrium conditions for the eligible users for which $\mu_g > 0$, which establishes our claim.

C.5 Proof of Lemma 5

We prove this claim by contradiction. Suppose for contradiction that $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$, where $x_{e,t}(\boldsymbol{\tau})$ denotes the equilibrium aggregate flow on edge e at period t under the CBCP scheme with toll $\boldsymbol{\tau}$. We use this relation to first show that the number of ineligible users on the express lane at period t, denoted as $x_{1,t}(\tilde{\boldsymbol{\tau}})_\mathcal{I}$ and $x_{1,t}(\boldsymbol{\tau})_\mathcal{I}$ for the two tolls must satisfy $x_{1,t}(\tilde{\boldsymbol{\tau}})_\mathcal{I} \leq x_{1,t}(\boldsymbol{\tau})_\mathcal{I}$, and then use this fact to derive our desired contradiction.

To establish that $x_{1,t}(\tilde{\boldsymbol{\tau}})_\mathcal{I} \leq x_{1,t}(\boldsymbol{\tau})_\mathcal{I}$, first note by the equilibrium condition for the ineligible users that a user in group $g \in \mathcal{G}_I$ takes the express lane at period t only when

$$v_{t,g} l_1(x_{1,t}(\tilde{\boldsymbol{\tau}})) + \tilde{\tau}_t \leq v_{t,g} l_2(x_{2,t}(\tilde{\boldsymbol{\tau}})).$$

Then, noting that $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$ and correspondingly that $x_{2,t}(\tilde{\boldsymbol{\tau}}) < x_{2,t}(\boldsymbol{\tau})$ as the user demand is fixed, consider an ineligible user group g that uses the express lane at equilibrium at period t under the toll $\tilde{\boldsymbol{\tau}}$. It then holds that

$$
\begin{aligned}
v_{t,g} l_1(x_{1,t}(\boldsymbol{\tau})) + \tau_t &\stackrel{(a)}{<} v_{t,g} l_1(x_{1,t}(\tilde{\boldsymbol{\tau}})) + \tilde{\tau}_t, \\
&\stackrel{(b)}{\leq} v_{t,g} l_2(x_{2,t}(\tilde{\boldsymbol{\tau}})), \\
&\stackrel{(c)}{<} v_{t,g} l_2(x_{2,t}(\boldsymbol{\tau})),
\end{aligned}
$$

where (a) follows as $\tilde{\tau}_t > \tau_t$ and by the assumption that $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$, (b) holds as the ineligible user group uses the express lane at period t under the toll $\tilde{\boldsymbol{\tau}}$, and (c) follows as $x_{2,t}(\tilde{\boldsymbol{\tau}}) < x_{2,t}(\boldsymbol{\tau})$. The above (strict) inequalities imply that if an ineligible user takes the express lane at period t under toll $\tilde{\boldsymbol{\tau}}$, then this user also takes the express lane under the toll $\boldsymbol{\tau}$. That is, $x_{1,t}(\tilde{\boldsymbol{\tau}})_\mathcal{I} \leq x_{1,t}(\boldsymbol{\tau})_\mathcal{I}$.

Next, let $x_{1,t}(\boldsymbol{\tau})_\mathcal{E}$ denote the equilibrium aggregate flow on edge e of eligible users under the CBCP scheme with toll $\boldsymbol{\tau}$. Then, the above derived relation for ineligible users, together with the fact that $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$, implies that $x_{1,t}(\tilde{\boldsymbol{\tau}})_\mathcal{E} > x_{1,t}(\boldsymbol{\tau})_\mathcal{E}$. We now show that this relation for the eligible users cannot hold true to derive the desired contradiction.

To see this, recall that under a set of tolls $\boldsymbol{\tau}$ eligible users use the express lane at the periods when the travel bang-per-buck ratio $\frac{v_{t,g}(l_2(x_{2,t}(\boldsymbol{\tau})) - l_1(x_{1,t}(\boldsymbol{\tau})))}{\tau_t}$ is the highest. Then, since the express lane flows satisfy $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$, the new ratio $\frac{v_{t,g}(l_2(x_{2,t}(\boldsymbol{\tau})) - l_1(x_{1,t}(\boldsymbol{\tau})))}{\tau_t} > \frac{v_{t,g}(l_2(x_{2,t}(\tilde{\boldsymbol{\tau}})) - l_1(x_{1,t}(\tilde{\boldsymbol{\tau}})))}{\tilde{\tau}_t}$. Further, the fact that $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$ implies that there are users from some group g such that $y_{1,t}^g(\tilde{\boldsymbol{\tau}}) > y_{1,t}^g(\boldsymbol{\tau})$, where $\mathbf{y}(\boldsymbol{\tau}), \mathbf{y}(\tilde{\boldsymbol{\tau}})$ represent the equilibrium flows corresponding to the tolls $\boldsymbol{\tau}, \tilde{\boldsymbol{\tau}}$, respectively. Then, by the budget

constraint of users in group g it must hold that there is some period t' at which the flow of users in group g on the express lane reduces under the toll $\tilde{\boldsymbol{\tau}}$, i.e., $y^g_{1,t'}(\tilde{\boldsymbol{\tau}}) < y^g_{1,t'}(\boldsymbol{\tau})$. However, for $\mathbf{y}(\tilde{\boldsymbol{\tau}})$ to be an equilibrium, it must hold that for all periods t' such that the flow of users in group g on the express lane reduces that the aggregate express lane flow must increase. Note if the aggregate express lane flow at one of the periods t' decreases under the toll $\tilde{\boldsymbol{\tau}}$, then users in group g have a profitable deviation at the new flow $\mathbf{y}(\tilde{\boldsymbol{\tau}})$, as they would achieve a lower travel cost when using the express lane at that period as compared to increasing their flow on the express lane at period t (given the change in their travel bang-per-buck ratios).

Next, since the flow of users in group g decrease at a subset of periods t' when the aggregate express lane flow is higher under the toll $\tilde{\boldsymbol{\tau}}$ (relative to that under the toll $\boldsymbol{\tau}$), there are other groups g' such that their flow is higher at the periods t'. By the above argument for group g, the flow $\mathbf{y}(\tilde{\boldsymbol{\tau}})$ can only be an equilibrium if for all groups g'' their express lane flow increases at one of the periods with a higher aggregate express lane flow. However, there are a subset of periods for which the total aggregate express lane flow increases under the toll $\tilde{\boldsymbol{\tau}}$ (which holds by our assumption that $x_{1,t}(\tilde{\boldsymbol{\tau}}) > x_{1,t}(\boldsymbol{\tau})$ at period t) and another set of periods for which the total aggregate express lane flow decreases (which holds by eligible users' budget constraints). Thus, it must be that under the toll $\tilde{\boldsymbol{\tau}}$ there is some user group such that their flow decreases at a period when the aggregate express lane flow decreases and increases at a period when the aggregate express lane flow increases. However, then the flow $\mathbf{y}(\tilde{\boldsymbol{\tau}})$ cannot be an equilibrium, which gives us our desired contradiction, proving our claim.

C.6   Proof of Proposition 1

Let the number of periods $T = 2$, where the toll on the express lane is equal in both periods, i.e., $\tau = \tau_1 = \tau_2$, and consider a setting with one eligible group (with total mass of one unit) with budget $B = 3\frac{\tau}{2}$ and one ineligible user group (with total mass of one unit). Further, let the eligible users have a time-invariant value of time, the travel time on the edges be such that $l_2(0.5) > l_1(1.5)$, and suppose that the ineligible users are such that they have a high value of time in the first period and a low value of time in the second period such that the ineligible users always use the express lane in period one and do not use the express lane in period two. Then, the equilibrium is such that only the eligible user group uses the express lane at period two and 0.5 units of the eligible user group (and the ineligible user group) use the express lane at period one. In particular, the following relation holds (with strict inequality):

$$\frac{l_2(1) - l_1(1)}{\tau} > \frac{l_2(0.5) - l_1(1.5)}{\tau}$$

From the above relation, it follows by continuity that the following relation must hold for some $\epsilon > 0$

$$\frac{l_2(1) - l_1(1)}{\tau(1 + \epsilon)} > \frac{l_2(0.5 + \epsilon) - l_1(1.5 - \epsilon)}{\tau}.$$

That is, the above relation implies that if the toll at period two is increased to $\tau(1 + \epsilon)$, the flow of the eligible users on the express lane reduces in the first period, which proves our claim.

C.7   Proof of Lemma 6

To prove this claim, we proceed by contradiction as in the proof of Lemma 5. In particular, we suppose that there is some period t such that $x_{1,t}(B) > x_{1,t}(\tilde{B})$, where $x_{e,t}(B)$ denotes the equilibrium aggregate flow on edge e at period t under the CBCP scheme where eligible users receive a budget B. Since the tolls for the two CBCP schemes are the same, we can show using a similar line of reasoning for ineligible users to that in the proof of Lemma 5 that when $x_{1,t}(B) > x_{1,t}(\tilde{B})$, it must hold that $x_{1,t}(B)_{\mathcal{I}} \leq x_{1,t}(\tilde{B})_{\mathcal{I}}$, where $x_{e,t}(B)_{\mathcal{I}}$ denotes the equilibrium aggregate flow of ineligible users on edge e at period t under the CBCP scheme where eligible users receive a budget B. As a result for $x_{1,t}(B) > x_{1,t}(\tilde{B})$ to hold, it must follow that $x_{1,t}(B)_{\mathcal{E}} > x_{1,t}(\tilde{B})_{\mathcal{E}}$, where $x_{e,t}(B)_{\mathcal{E}}$ denotes the equilibrium aggregate flow of eligible users on edge e at period t under the CBCP scheme where eligible users receive a budget B. We now show that the relation that $x_{1,t}(B)_{\mathcal{E}} > x_{1,t}(\tilde{B})_{\mathcal{E}}$ cannot hold true to derive our desired contradiction.

To see this, recall that, given a budget B and express lane tolls $\boldsymbol{\tau}$, eligible users use the express lane at the periods when the travel bang-per-buck ratio $\frac{v_{t,g}(l_2(x_{2,t}(B)) - l_1(x_{1,t}(B)))}{\tau_t}$ is the highest. Then, since the express

lane flows satisfy $x_{1,t}(B) > x_{1,t}(\tilde{B})$, the new ratio $\frac{v_{t,g}(l_2(x_{2,t}(B))-l_1(x_{1,t}(B)))}{\tau_t} > \frac{v_{t,g}(l_2(x_{2,t}(\tilde{B}))-l_1(x_{1,t}(\tilde{B})))}{\tau_t}$. Further, the fact that $x_{1,t}(B) > x_{1,t}(\tilde{B})$ implies that there are users from some group g such that $y_{1,t}^g(B) > y_{1,t}^g(\tilde{B})$, where $\mathbf{y}(B)$, $\mathbf{y}(\tilde{B})$ represent the equilibrium flows corresponding to the budgets $B, \tilde{B}$, respectively. Then, by the budget constraint of users in group g it must hold that there is some period t' at which the flow of users in group g on the express lane is lower under the budget B, i.e., $y_{1,t'}^g(\tilde{B}) > y_{1,t'}^g(B)$. However, for $\mathbf{y}(B)$ to be an equilibrium, it must hold that for all periods t' such that the flow of users in group g on the express lane reduces that the aggregate express lane flow must increase. Note if the aggregate express lane flow at one of the periods t' decreases under the budget B (relative to that under the budget $\tilde{B}$), then users in group g have a profitable deviation at the flow $\mathbf{y}(B)$, as they would achieve a lower travel cost when using the express lane at that period as compared to increasing their flow on the express lane at period t (given the change in their travel bang-per-buck ratios).

Next, since the flow of users in group g decrease at a subset of periods t' when the aggregate express lane flow is higher under the budget B (relative to that under the budget $\tilde{B}$), there are other groups g' such that their flow is higher at the periods t'. By the above argument for group g, the flow $\mathbf{y}(B)$ can only be an equilibrium if for all groups g'' their express lane flow increases at one of the periods with a higher aggregate express lane flow. However, there are a subset of periods for which the total aggregate express lane flow increases under the toll B (which holds by our assumption that $x_{1,t}(B) > x_{1,t}(\tilde{B})$ at period t) and another set of periods for which the aggregate express lane flow decreases (which holds by eligible users' budget constraints). Thus, it must be that under the budget B there is some user group such that their flow decreases at a period when the aggregate express lane flow decreases and increases at a period when the aggregate express lane flow increases. However, then the flow $\mathbf{y}(B)$ cannot be an equilibrium, which gives us our desired contradiction, proving our claim.

C.8 Proof of Proposition 2

Let the number of periods $T = 1$ and suppose that all users are eligible with a total demand of one unit. Further, suppose that the toll on the express lane is τ and the budget given to eligible users is $B = \frac{\tau}{2}$. In addition, consider a travel time function where $l_1(0.5) = 0.5$, $l_1(0.75) = 1.5$, $l_2(0.5) = 2$, and $l_2(0.25) = 1.99$. Then, since fractional flows are allowed, the resulting equilibrium flow pattern will be such that 0.5 units of users use the express lane and 0.5 units of users use the general-purpose lane, resulting in a total travel cost of 1.25 (as $0.5 \times 0.5 + 2 \times 0.5 = 1.25$) for each user. On the other hand, if the budget were increased to $\tilde{B} = \frac{3\tau}{4}$, then the resulting equilibrium flow pattern will be such that 0.75 units of users use the express lane and 0.25 units of users use the general-purpose lane, resulting in a total travel cost of 1.6225 (as $1.5 \times 0.75 + 1.99 \times 0.25 = 1.6225$) for each user. These relations establish that an increase in the budget of eligible users can in fact increase their travel costs even if their values of time are time-invariant, which proves our claim.

C.9 Proof of Lemma 7

To prove this claim, we leverage Berge's theorem, which establishes continuity relations of the optimal solutions of parametric optimization problems in their parameters.

Theorem 3 (Berge's Theorem [46]). *Consider the parametric constrained maximization problem*

$$\max F(z, \theta), \ s.t. \ z \in A(\theta).$$

Let $Z(\theta)$ be the set of solutions of this problem with parameter θ. If

(a) *F is a continuous function in (z, θ),*

(b) *$\theta \implies A(\theta)$ is lower semi-continuous, and*

(c) *there exists for each θ a set $B(\theta) \subseteq A(\theta)$ such that $Z(\theta) \subseteq B(\theta)$, $\sup\{F(z, \theta) : z \in B(\theta)\} = \sup\{F(z, \theta) : z \in A(\theta)\}$, and $\theta \implies B(\theta)$ is an upper semi-continuous and locally bounded correspondence.*

Then:

1. *$Z(\theta)$ is non-empty for all θ, and $\theta \implies Z(\theta)$ is an upper semi-continuous and locally bounded correspondence.*

2. *If $Z(\theta)$ is a singleton set, i.e., $Z(\theta) = \{z(\theta)\}$ for all θ in an open set of parameter values, then $z(\theta)$ is a continuous function over that set of parameter values.*

We now show that the convex Program (4a)-(4c) satisfies conditions (a), (b), and (c) of Berge's theorem to derive the desired continuity relations. To this end, first note that condition (a) of Theorem 3 clearly holds as the Objective (4a) is continuous in $(\mathbf{y}, (\boldsymbol{\tau}, B))$. In particular, the Objective (4a) is independent of the budget B, depends linearly on the toll $\boldsymbol{\tau}$, and is continuous in the edge flows by the continuity of the travel time function. Next, let $\mathcal{F}(\boldsymbol{\tau}, B)$ denote the set of all feasible flows satisfying Constraints (4b)-(4c). We now establish that conditions (b) and (c) of Theorem 3 hold by establishing that the correspondence $(\boldsymbol{\tau}, B) \implies \mathcal{F}(\boldsymbol{\tau}, B)$ is (i) locally bounded, (ii) upper semi-continuous, and (iii) lower semi-continuous.

Local Boundedness of $\mathcal{F}(\boldsymbol{\tau}, B)$: By the definition of a locally bounded correspondence (Definition 4), we show for every $(\boldsymbol{\tau}, B) \in \mathcal{F}_U$ that there is an $\epsilon(\boldsymbol{\tau}, B) > 0$ and a bounded set $\mathcal{Y}(\bar{\boldsymbol{\tau}}, \bar{B})$ is a subset of this bounded set for all $(\bar{\boldsymbol{\tau}}, \bar{B})$ that lie within an ϵ ball of $(\boldsymbol{\tau}, B)$. To this end, fix $(\boldsymbol{\tau}, B)$ and let $\epsilon = \frac{1}{3}\min_{t=1,\ldots,T}\tau_t$. Next, observe that for all $(\bar{\boldsymbol{\tau}}, \bar{B})$ that lie within an ϵ ball of $(\boldsymbol{\tau}, B)$ that any feasible flow $\mathbf{y} \in \mathcal{F}(\bar{\boldsymbol{\tau}}, \bar{B})$ must satisfy $\bar{\tau}_1^t y_{1,t}^g \leq \bar{B}$ by the budget Constraint (4c). This relation implies that $y_{1,t}^g \leq \frac{B+\epsilon}{\bar{\tau}_1^t}$ from which it follows by the definition of ϵ that $y_{1,t}^g \leq \frac{B+\epsilon}{2\epsilon}$, which provides a uniform bound on $\mathbf{y} \in \mathcal{F}(\bar{\boldsymbol{\tau}}, \bar{B})$. This establishes that the correspondence $\mathcal{F}(\boldsymbol{\tau}, B)$ is locally bounded.

Upper Semi-Continuity of $\mathcal{F}(\boldsymbol{\tau}, B)$: To show upper semi-continuity of $\mathcal{F}(\boldsymbol{\tau}, B)$, we consider a sequence $(\{\mathbf{y}^n, \boldsymbol{\tau}^n, B^n\})$ where $\mathbf{y}^n \in \mathcal{F}(\boldsymbol{\tau}^n, B^n)$ for all n with a limit $(\mathbf{y}, \boldsymbol{\tau}, B)$. As $\mathbf{y}^n \geq \mathbf{0}$ for all n and since the positive orthant is closed, it follows that $\mathbf{y} \geq 0$. Next, since $\sum_{t=1}^T \tau_t^n y_{1,t}^{gn} \leq B^n$ for all n for all user groups g it follows by the continuity of the dot product that $\sum_{t=1}^T \tau_t y_{1,t}^g \leq B$. Finally, since $y_{1,t}^{gn} + y_{2,t}^{gn} = 1$ for all n it also follows that $y_{1,t}^g + y_{2,t}^g = 1$ for all eligible user groups $g \in \mathcal{G}_E$ and periods t. Hence, we have established that $\mathbf{y} \in \mathcal{F}(\boldsymbol{\tau}, B)$, which proves upper semi-continuity.

Lower Semi-Continuity of $\mathcal{F}(\boldsymbol{\tau}, B)$: To establish lower semi-continuity of $\mathcal{F}(\boldsymbol{\tau}, B)$, we show that for each sequence $(\boldsymbol{\tau}^n, B^n)$ with limit $(\boldsymbol{\tau}, B)$ and a point $\mathbf{y} \in \mathcal{F}(\boldsymbol{\tau}, B)$ that there is a sequence $\mathbf{y}^n$ with limit $\mathbf{y}$ such that $\mathbf{y}^n \in \mathcal{F}(\boldsymbol{\tau}^n, B^n)$ for all n. To show this, first consider the case when $B = 0$. In this setting, by the positivity of the tolls it follows that $y_{1,t}^g = 0$ and $y_{2,t}^g = 1$ for all eligible user groups g. Thus, the sequence where $y_{1,t}^{gn} = 0$ and $y_{2,t}^{gn} = 1$ holds for all eligible user groups is feasible for all n. Analogously, if the flow $\mathbf{y}_1^g = \mathbf{0}$ (where $\mathbf{y}_1 = (y_{1,t}^g : t \in [T])$), i.e., there is no flow of eligible users in group g on the express lane, then setting $\mathbf{y}_1^{gn} = \mathbf{0}$ results in a feasible sequence. Thus, consider the setting when for some user group g $\mathbf{y}_1^g \neq \mathbf{0}$, the budget $B > 0$, and consider n large enough such that $\sum_{t=1}^T \tau_t^n y_{1,t}^g$ is uniformly bounded away from zero. Next, let $y_{1,t}^{gn} = \min\left\{\frac{B^n}{B}\frac{\sum_{t=1}^T \tau_t y_{1,t}^g}{\sum_{t=1}^T \tau_t^n y_{1,t}^g}y_{1,t}^g, 1\right\}$ and $y_{2,t}^{gn} = 1 - y_{1,t}^{gn}$. Then, it is clear that $y_{1,t}^{gn} \geq 0$ and

$y_{1,t}^{gn} + y_{2,t}^{gn} = 1$. Furthermore, the budget constraint is also satisfied as

$$\sum_{t=1}^{T} \tau_t^n y_{1,t}^{gn} = \sum_{t=1}^{T} \tau_t^n \min \left\{ \frac{B^n}{B} \frac{\sum_{t=1}^{T} \tau_t y_{1,t}^g}{\sum_{t=1}^{T} \tau_t^n y_{1,t}^g} y_{1,t}^g, 1 \right\},$$

$$\leq \min \left\{ \frac{B^n}{B} \frac{\sum_{t=1}^{T} \tau_t y_{1,t}^g}{\sum_{t=1}^{T} \tau_t^n y_{1,t}^g} \sum_{t=1}^{T} \tau_t^n y_{1,t}^g, \sum_{t=1}^{T} \tau_t^n \right\},$$

$$\leq \min \left\{ B^n \frac{\sum_{t=1}^{T} \tau_t y_{1,t}^g}{B}, \sum_{t=1}^{T} \tau_t^n \right\},$$

$$\leq \min \left\{ B^n, \sum_{t=1}^{T} \tau_t^n \right\},$$

$$\leq B^n.$$

Thus, for all eligible user groups (depending on whether $\mathbf{y}_1^g = \mathbf{0}$ or $\mathbf{y}_1^g \neq \mathbf{0}$) we can construct a sequence of $y_{e,t}^{gn}$ such that $\mathbf{y}^n \in \mathcal{F}(\boldsymbol{\tau}^n, B^n)$ for all n, which establishes lower semi-continuity.

Having established that $(\boldsymbol{\tau}, B) \implies \mathcal{F}(\boldsymbol{\tau}, B)$ is (i) locally bounded, (ii) upper semi-continuous, and (iii) lower semi-continuous, the upper semi-continuity and local boundedness of the correspondence $\mathcal{Y}(\boldsymbol{\tau}, B)$ follows from a direct application of Berge's theorem (Theorem 3). Furthermore, noting from Lemma 4 that the aggregate edge flows are unique, it follows from Berge's theorem that the aggregate edge flows $\mathbf{x}(\boldsymbol{\tau}, B)$ are continuous in $(\boldsymbol{\tau}, B)$, which proves our claim.

D   Additional Experimental Results

In this section, we present additional detailed metrics and graphics corresponding to the experimental results.

D.1   PeMS Data

Table 2: US 101 Express Lanes PeMS Utilization Data (September, 2022)

| | | Station ID (Postmile location) | | | |
| | | 400388 | 405673 | 404532 | |
	Hour	(405.4)	(406.9)	(408.4)	Mean
% Flow on express	7 - 8 AM	14%	14%	15%	14%
	8 - 9 AM	16%	15%	15%	15%
	9 - 10 AM	13%	12%	13%	13%
	7 - 10 AM	14%	14%	14%	14%
Mean mph express lanes	7 - 8 AM	74	70	67	70
	8 - 9 AM	66	63	65	65
	9 - 10 AM	67	69	65	67
	7 - 10 AM	69	67	66	67
Mean mph GP lanes	7 - 8 AM	51	44	51	49
	8 - 9 AM	40	36	48	41
	9 - 10 AM	43	44	50	46
	7 - 10 AM	45	42	50	45
Mean travel time savings	7 - 8 AM	32%	37%	24%	31%
	8 - 9 AM	40%	43%	26%	36%
	9 - 10 AM	36%	35%	23%	32%
	7 - 10 AM	36%	38%	24%	33%

D.2 Experimental results

Table 3 includes the optimal CBCP parameter (toll and budget levels) as well as the equilibrium percentage of users on the express lane by group, the average travel times, average travel costs by group, and the total toll revenue for 18 different Pareto weighted schemes for the objective function in Section 7.3. Furthermore, Figure 6 displays the distributions of the optimal a) CBCP parameters, b) express lane usage, c) average travel times, and d) average social costs and total toll revenue with respect to the value of the Pareto weight for eligible users, holding the remaining two weights (λ_I, λ_R) equal to 1.

Table 3: Optimal CBCP and Equilibrium Edge Flows and Travel Times for Various Pareto Weighted Schemes

Weights			Optimal CBCP ($)		Express lane usage (%)			Average TTs (minutes)		Total travel costs ($) (% diff from (1,1,1))		Total toll revenue ($)
λ_e	λ_i	λ_t	τ	B	Overall	Eligible	Ineligible	Express	GP	Eligible	Ineligible	(% diff from (1,1,1))
1	0	0	19	90	19	95	3	22.1	30.3	17738 (-27%)	864766 (1%)	21337 (-76%)
0	1	0	0	0	25	60	18	28.2	28.2	22210 (-9%)	808169 (-6%)	0 (-100%)
0	0	1	15	0	16	0	19	19.4	31.4	24664 (2%)	868444 (1%)	94098 (6%)
1	1	1	13	0	17	0	21	20.3	30.9	24282 (0%)	856500 (0%)	89097 (0%)
2	1	1	12	0	18	0	21	20.9	30.7	24133 (-1%)	852576 (0%)	84963 (-5%)
3	1	1	12	0	18	0	21	20.9	30.7	24133 (-1%)	852576 (0%)	84963 (-5%)
4	1	1	12	0	18	0	21	20.9	30.7	24133 (-1%)	852576 (0%)	84963 (-5%)
5	1	1	11	0	18	0	22	21.5	30.5	23983 (-1%)	848654 (-1%)	80390 (-10%)
6	1	1	11	0	18	0	22	21.5	30.5	23983 (-1%)	848654 (-1%)	80390 (-10%)
7	1	1	11	0	18	0	22	21.5	30.5	23983 (-1%)	848654 (-1%)	80390 (-10%)
8	1	1	10	0	19	0	23	22.1	30.3	23831 (-2%)	844763 (-1%)	75391 (-15%)
9	1	1	10	0	19	0	23	22.1	30.3	23831 (-2%)	844763 (-1%)	75391 (-15%)
10	1	1	10	0	19	0	23	22.1	30.3	23831 (-2%)	844763 (-1%)	75391 (-15%)
11	1	1	10	15	19	30	17	22.3	30.2	21900 (-10%)	846033 (-1%)	56060 (-37%)
12	1	1	11	45	19	82	7	22.6	30.1	18850 (-22%)	848827 (-1%)	24101 (-73%)
13	1	1	12	50	19	83	6	22.4	30.2	18611 (-23%)	851683 (-1%)	23890 (-73%)
14	1	1	13	55	19	85	6	22.2	30.3	18424 (-24%)	854178 (0%)	23855 (-73%)
15	1	1	13	55	19	85	6	22.2	30.3	18424 (-24%)	854178 (0%)	23855 (-73%)

(a) Toll and budget

(b) Share of users using express lane

(c) Average travel times
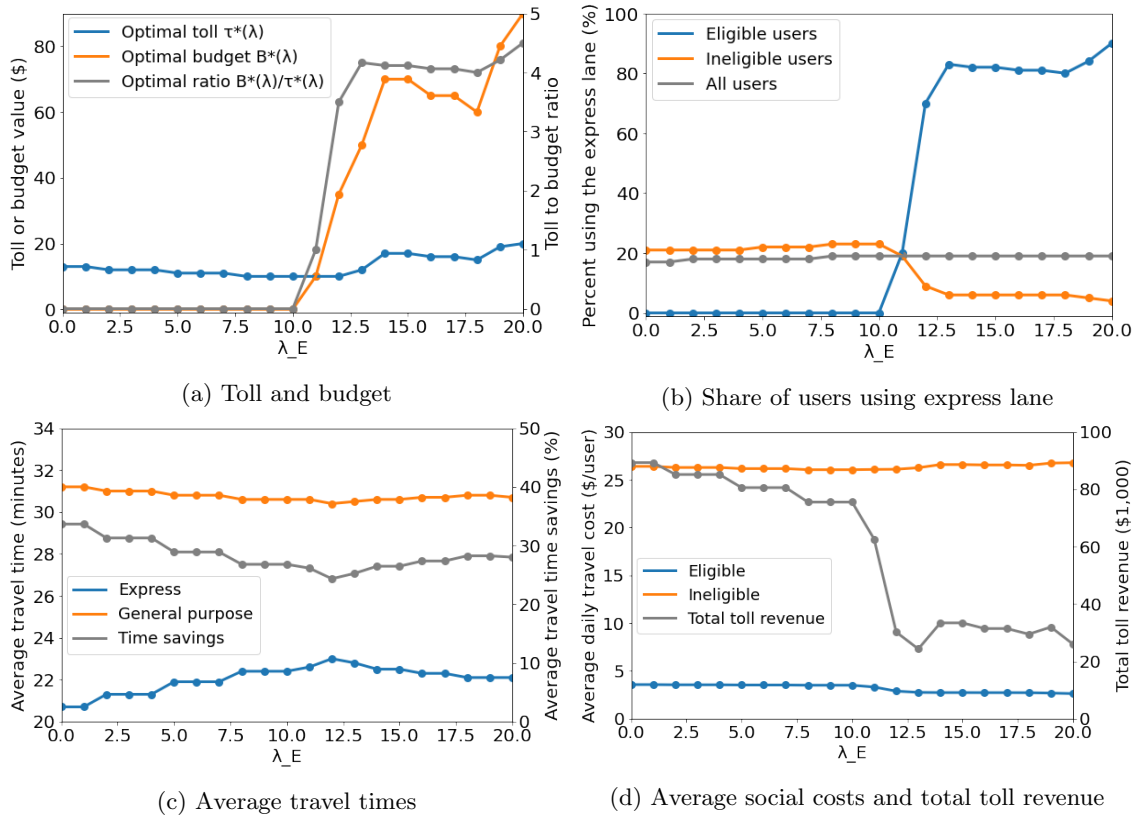
(d) Average social costs and total toll revenue

Figure 6: Distributions of the optimal a) CBCP parameters, b) express lane usage, c) average travel times, and d) average social costs and total toll revenue with respect to the value of the Pareto weight for eligible users, holding the remaining two weights (λ_I, λ_R) equal to 1.