2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Incentivizing Behavior in Transportation Networks with Non-Rational Drivers and 3rd-party Economic Agents

Colton Hill and Philip N. Brown

***Abstract*— A social planner who wishes to influence human decision-making must shape incentives to account for non-rational decision-making biases among the population to be influenced. However, the planner does not operate in a vacuum; societies are comprised of many heterogeneous economic actors whose self-interested behavior may hamper the social planner's ability to achieve their goals. For instance, in a transportation network which is subject to road tolls, a 3rd-party economic agent (whom we call the arbitrageur) may launch a service to provide users with information which helps them optimize their toll pricing and congestion experiences. What are the effects of such an arbitrageur on the social planner's incentive design problem? Do there still exist behavior-optimizing incentive schemes in the presence of an arbitrageur? In this work, our contributions are a formal model of this scenario and analytical derivations of game-theoretic equilibria for a simple transportation network.**

I. Introduction

In today's world of smart infrastructure, social media, and misinformation, large-scale attempts to influence behavior in society are ubiquitous [1]. Many entities attempt to wield influence over social behavior using diverse tools such as tax policy, misinformation-for-hire hacker groups, and nudge-theoretic techniques [2]. How effective can these behavior-change approaches truly be in a society comprising tens of thousands of influencers and billions of influencees?

A first challenge in modeling social influence schemes is that human social behavior is widely understood to be sub-rational. That is, human decisionmakers (such as drivers in a transportation network) are not expected-utility maximizers, but rather are subject to behavioral biases which shape their decisions [3]. Hence, a social planner wishing to influence behavior may need to shape the offered incentives to account for the behavioral biases of the human population.

A second complicating factor in deploying social influence schemes is that models must take into account the potential role of 3rd-party economic actors. The social planner and the population to be influenced do not operate in a vacuum, and some types of incentive schemes may provide self-interested 3rd parties with opportunities to profitably interfere, potentially undermining the effectiveness of the applied incentives.

Congestion control via road tolls in highway networks is a common testbed application area for studies in this domain. In isolation, the incentive (pricing) problem in transportation networks is well studied; the main goal is to assign prices to roads in a transportation network in such a way that when all drivers choose network routes in a self-interested way (i.e., minimizing a sum of delay and monetary cost), the overall congestion on the network is minimized [4]–[6]. In addition, much is known about population behavior in transportation networks in which the driver population makes selfish *biased* decisions of various kinds [7]–[9].

However, to our knowledge, all existing research on the transportation problem has treated the social planner and the drivers as the only agents in the system. What effects can 3rd-party economic actors have when a social planner attempts to shape an incentive mechanism to a particular set of behavioral biases? For example, suppose the driver population is known to irrationally over-weight the importance of tolls, and the social planner deploys a monetary incentive scheme which optimizes network routing given these biases. If the driver population reacts to the incentive scheme in accordance with their usual behavioral biases, then the social state which results is

- socially optimal,
- a biased Nash equilibrium (i.e., the drivers are individually satisfied in light of their behavior biases), but
- *not* a Nash equilibrium from the standpoint of expected utility maximization theory (i.e., some drivers could switch strategies and realize a *real* gain, despite their biases telling them otherwise).

In this scenario, there is "money left on the table," potentially presenting a clever 3rd-party with an *arbitrage opportunity*: Some system users could gain by switching strategies, so an arbitrageur could conceivably find a way to capture some of the utility surplus (for instance, by selling users a subscription to an "incentive optimization service"), which would effectively shift the social state to some heterogeneous Nash equilibrium that is *not* a social optimum.

In this paper, we formally pose a general model of this type of interaction, and we fully characterize the equilibria resulting from an arbitrageur's actions for fixed tolls on a 2-link Pigou network.

II. Contribution: An Economically-Embedded Routing Problem

Our first contribution is to formally pose a complete model of influence by a 3rd-party informational arbitrageur in transportation networks. This model allows for the driver population to have very general behavioral biases and for the arbitrageur to shift traffic flow arbitrarily. In Section III, we instantiate this model on a simple network for a specific type of behavioral bias and derive preliminary results.

*This material is based upon work supported by the National Science Foundation under award number ECCS-2013779.

The authors are with the University of Colorado at Colorado Springs, CO 80918, USA. {chill13,pbrown2}@uccs.edu

A. Network Routing Problem

Consider a network routing problem in which a unit mass of traffic needs to be routed across a network (V, E), which consists of a vertex set V and edge set $E \subseteq (V \times V)$ with a given source $s \in V$ and destination $t \in V$. We write $\mathcal{P} \subset 2^E$ to denote the set of *paths* available to traffic, where each path $p \in \mathcal{P}$ consists of a set of edges connecting s to t. We write $f_p \geq 0$ to denote the mass of traffic using path p. A *feasible flow* $f \in \mathbb{R}^{|\mathcal{P}|}$ is an assignment of traffic to various paths such that $\sum_{p \in \mathcal{P}} f_p = r$.

Given a flow f, the flow on edge e is given by $f_e = \sum_{p: e \in p} f_p$. To characterize transit delay as a function of traffic flow, each edge $e \in E$ is associated with a specific latency function $\ell_e : [0,1] \to [0, \infty)$. We adopt the standard assumptions that latency functions are nondecreasing, continuously differentiable, and convex. The system-level cost of a flow f is measured by the *total latency*, given by

$$\mathcal{L}(f) = \sum_{e \in E} f_e \cdot \ell_e(f_e) = \sum_{p \in \mathcal{P}} f_p \cdot \ell_p(f), \tag{1}$$

where $\ell_p(f) = \sum_{e \in p} \ell_e(f_e)$ denotes the latency on path p.

Our core goal is to understand the extent to which it is possible to influence social behavior using monetary incentives *in an economically-embedded routing problem*; that is, when other economic entities exist to capture surpluses potentially generated by the incentive schemes. Accordingly, we introduce the following game formulation to allow for modeling a wide variety of individual user biases in decision-making. We use a non-atomic game in which the traffic models a large population of drivers; note that we view each driver as infinitesimally small and having a negligible individual impact on traffic congestion. In order to influence driver behavior, the social planner assigns each edge $e \in E$ a (possibly flow-dependent) tolling function $\tau_e : [0,1] \to \mathbb{R}^+$. We assume that drivers generally wish to avoid some combination of delay and tolls, but we do not prescribe the specific biases which drive behavior.

B. Behavioral Biases

Much prior work has assumed that users are quasilinear in their preferences; that is, individual user costs are a convex combination of latency and tax [10]. However, in this project, we place few specific restrictions on how tax- and latency-aversion may manifest. To model a wide variety of heterogeneous behavioral biases, we model the users as the interval $[0,1]$; each user $x \in [0,1]$ has a cost function $c^x(l,t)$ which models the subjective cost experienced by user x when considering an edge with latency l and tax t. To model cost aversion, we assume that for every x, $c^x(l,t)$ is nondecreasing in both l and t. One special bias function that we frequently consider is the *neutral* bias function $c^{\mathrm{n}}(l,t) = l + t$; we take c^{n} to represent the cost function employed by a completely rational cost-minimizer. As a reference point, we consider the neutral bias function to be an objective measure of drivers' experienced costs.

Given a flow f, the cost that user $x \in [0,1]$ experiences for using path $\tilde{p} \in \mathcal{P}$ is of the form

$$J^x(f;\tau) = \sum_{e \in \tilde{p}} c^x\left(\ell_e(f_e), \tau_e(f_e)\right). \tag{2}$$

To denote the neutral cost on path p, we write

$$J_p^{\mathrm{n}}(f;\tau) = \sum_{e \in p} c^{\mathrm{n}}\left(\ell_e(f_e), \tau_e(f_e)\right). \tag{3}$$

This is a sufficiently general framework to allow for a generalization of many previously-considered behavioral biases, including heterogeneous tax sensitivity [11] and prospect theory [12], [13]. We assume that each driver prefers the lowest-cost path from the available source-destination paths; despite the fact that drivers may not make their decision based solely on a sum of time and money, they still act in a "selfish" manner in light of their individual behavioral biases.

Let $F : [0,1] \to \mathcal{P}$ be a function which describes the specific path choices of each driver $x \in [0,1]$. Note that each F uniquely determines a network flow in which $f_p := \{x \in [0,1] : F(x) = p\}$. In the remainder of the paper, when a network flow f and action specification F appear together, we take f to be the network flow uniquely determined by F. We call F a *Nash equilibrium* if every driver $x \in [0,1]$ is using a minimum-cost path:

$$F(x) \in \operatorname*{arg\,min}_{p \in \mathcal{P}} J^x(f;\tau). \tag{4}$$

It is well-known that a Nash equilibrium exists for any non-atomic game of the above form [14]. Note that it will often be convenient to refer to a network flow f as a *Nash flow*; when we do so, it is to be understood that f is the network flow which is uniquely determined by some Nash equilibrium F.

C. Arbitrage

In a real socio-technical system, the economic context involves many ancillary decision-makers other than the system users and the infrastructure manager. In particular, in this paper we focus on the role played by a profit-seeking 3rd-party called an *arbitrageur* which develops a way to profit from the sub-rationality of drivers by informing them of lower-cost alternatives (of which the users were previously unaware due to their behavioral biases).

The high-level setup is this: in a biased Nash equilibrium F_τ^{b} under the influence of tolls τ, some users will generally be selecting routes with a relatively high neutral cost as measured by the neutral bias function $c^{\mathrm{n}}(l,t) = l + t$. The arbitrageur creates an information service that offers a subset of these inefficiently-routing users a route recommendation which would lower their neutral cost. If a user subscribes to the information service, the arbitrageur recommends a new route with a lower neutral cost; we assume that the user adopts this recommendation and switches to the recommended route.

In the biased Nash equilibrium F_τ^{b} (with associated network flow f^{b}), suppose that the arbitrageur informs a subset of the users of path p of the existence of lower-cost path q.

Formally, the arbitrageur informs a subset of $\{x \in [0,1] : F^{\mathrm{b}}_\tau(x) = p\}$ of mass $M_{pq} \geq 0$ and recommends they switch to path q; we write $M := (M_{pq})_{(p,q)\in\mathcal{P}\times\mathcal{P}}$. Once all of the informed users have switched paths, the remaining mass of uninformed users re-routes in a new *informed* Nash equilibrium $F^{\mathrm{i}}_{M,\tau}$.

An $F^{\mathrm{i}}_{M,\tau}$ as described always exists, since it can be viewed as a Nash equilibrium for a population in which every informed user x^{i} has fixed actions, and every uninformed user's bias is unchanged from the original problem.

The arbitrageur devises a way to capture a proportion of the neutral cost savings obtained by the informed traffic (e.g., by selling the information for a subscription fee or by monetizing the information using sponsored search); for simplicity, we use this cost savings as a proxy for the arbitrageur's profits. Denote the cost savings experienced by informed drivers who have switched from path p to path q by

$$\Delta_{pq}(M,\tau) := J^{\mathrm{n}}_p\left(f^{\mathrm{i}};\tau\right) - J^{\mathrm{n}}_q\left(f^{\mathrm{i}};\tau\right). \tag{5}$$

The arbitrageur's revenue (i.e., the sum of all driver individual cost savings) is thus given by

$$R(M;\tau) = \sum_{(p,q)\in\mathcal{P}\times\mathcal{P}} M_{pq}\Delta_{pq}(M,\tau). \tag{6}$$

D. The Infrastructure Manager's Problem

Now, the infrastructure manager wishes to levy tolls τ on the network to minimize the total latency (1); that is, the infrastructure manager's objective is to minimize a cost function of the form

$$\mathcal{L}(M,\tau) := \mathcal{L}(f^i(M;\tau)). \tag{7}$$

If the infrastructure manager is aware of the presence of the arbitrageur, this cost function induces a *Stackelberg game*; the arbitrageur's optimal (contingent) choice is

$$M^*(\tau) := \arg\max_M R(M;\tau), \tag{8}$$

and the infrastructure manager wishes to solve the problem

$$\tau^* = \arg\min_\tau \mathcal{L}(M^*(\tau),\tau). \tag{9}$$

Note that in general, τ^* may be quite a complex object due to the fact that it is a Stackelberg strategy of a two-layer game: first the infrastructure manager (Stackelberg leader) selects a set of tolls τ, then the arbitrageur (Stackelberg follower) selects a set of users M to inform, and finally the uninformed user population establishes a Nash equilibrium given the higher-level choices of τ and M.

III. Arbitrage on a Simple Network

To provide initial insights into the effects of arbitrage as described in Section II, we study an instance of the model on a simple network:

- Section III-A describes the specific network,
- Section III-B presents fixed and marginal-cost tolls,
- Section III-C derives the Nash flows without arbitrage,
- Section III-D derives the effects of arbitrage with fixed tolls, and
- Section III-E does the same with marginal-cost tolls.

A. Network

To illustrate an instance of this model and show that an arbitrageur may significantly change the strategic environment, consider the simple 2-link network depicted in Figure 1 (this is the classic Pigou network used to show many salient facts about selfish routing problems [15]).

Our network has two parallel links; $\ell_1(f_1) = f_1$ and $\ell_2(f_2) = 1$. The total latency (1) is minimized when the traffic is split evenly between the two links, which we denote $f^* = (1/2, 1/2)$, giving an optimal total latency of $\mathcal{L}(f^*) = 3/4$. However, the un-influenced (zero-toll) Nash flows have $f_1 = 1$ and $f_2 = 0$, with a total latency of 1. To improve network congestion, the infrastructure manager levies a tolling function $\tau_1(f_1)$ on the congestible upper link, hoping to cause some drivers to deviate to the constant-latency lower link.

B. Tolls

Two classical tolling approaches are *fixed tolls* [10] which are simply constant functions of flow, and *marginal-cost tolls* [16] which take the form $\tau^{\mathrm{mc}}_i(f_i) = f_i\ell'_i(f_i)$. In selecting between the two, an infrastructure manager must choose between the simplicity and predictability of fixed tolls and the robustness of marginal-cost tolls [4]. In this paper, we perform a preliminary comparison of the performance of these two classes of tolls in the presence of a 3rd-party arbitrageur. Note that in all cases, the infrastructure manager levies a toll only on link 1; i.e., $\tau_2(f_2) = 0$.

C. Nash Flows Without Arbitrage

In this paper, we consider populations which are broadly biased toward toll-aversion; each driver $x \in [0,1]$ has a bias function of $c^x(l,t) = l + Sxt$ for some fixed $S > 0$. Note that if $S > 1$, the population contains some drivers who underestimate the true cost of tolls (i.e., for some $x \in [0,1]$, it holds that $Sx < 1$), and some who overestimate the cost of tolls. When $S > 2$, the population is toll-averse "on average": the median driver $x = 1/2$ over-weights the cost of tolls by a factor of $S/2 > 1$. On the other hand, when $S < 2$, the population is latency-averse "on average," since the median driver $x = 1/2$ under-weights the cost of tolls by a factor of $S/2 < 1$.

Under this particular bias function, it is possible to compute drivers' equilibrium responses to tolls:

Proposition 3.1: When no arbitrageur is present, the biased Nash flow for the considered user population with $S > 0$ is $(f^{\mathrm{b}}(\tau) := (f^{\mathrm{b}}_1, 1 - f^{\mathrm{b}}_1)$ where

$$f^{\mathrm{b}}_1 := \max\{0, \phi\} \tag{10}$$

and ϕ is the unique positive solution to:

$$\phi + S\phi\tau_1(\phi) = 1. \tag{11}$$

Proof: Let f^{b}_1 be an arbitrary Nash flow. If user x is on link 1 with respect to f^{b}_1, then $f^{\mathrm{b}}_1 + Sx\tau_1 \leq 1$, that is, $x \leq \frac{1-f^{\mathrm{b}}_1}{S\tau_1}$. Similarly, if user x is on link 2, with respect to f^{b}_1, then $x \geq \frac{1-f^{\mathrm{b}}_1}{S\tau_1}$. Hence, lower-sensitivity users will go

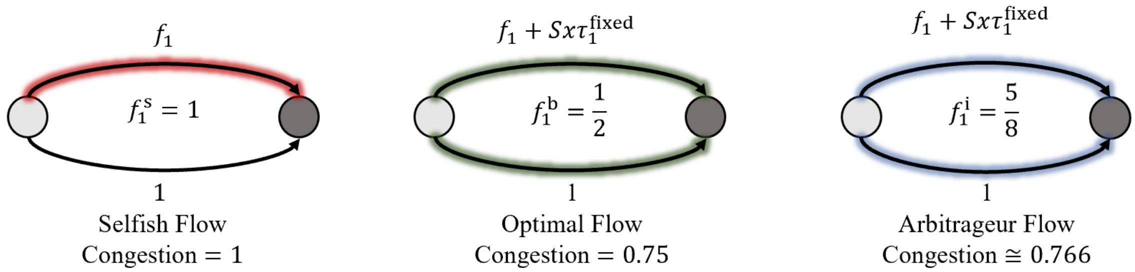

Fig. 1: An example of the problems that arise in the presence of an arbitrageur in a simple Pigou traffic network. Here, $S = 4$, $\tau_1^{\mathrm{fixed}} = 1/4$, and an arbitrageur informs revenue-maximizing mass $M^*(\tau) = 1/4$ of traffic to move from link 1 to link 2. If an infrastructure manager decides to assess no tolls to a traffic network (the left network), the resulting congestion can be far from optimal. Hence, the planner may choose to charge tolls to certain roads in the network in order to obtain improved (optimal) total latency (the middle network). However, a self-interested third party may seek to gain revenue by informing drivers of less costly roads—often at the expense of once-again suboptimal total latency.

on link 1, while higher-sensitivity users will occupy link 2. Now, there is at least one user on link 1, otherwise, for user 0, $f_1^{\mathrm{b}} + S0\tau_1 = f_1^{\mathrm{b}} > 1$, contradicting $f_1^{\mathrm{b}} \leq 1$. Likewise, there is at least one user on link 2, otherwise, for user 1, $f_1^{\mathrm{b}} + S\tau_1 < 1$, i.e. $S\tau_1 < 0$, a contradiction since $S > 0$ and $\tau \geq 0$. Thus, we have that at least one user is on link 1, and the most sensitive user on link 1 is indifferent between link 1 and link 2. Now, since all users are represented as a continuum in $[0, 1]$, the most sensitive user on link one (say, $\bar{x}$) is indifferent, and equal to the flow on link 1. Hence, there is a unique indifferent user $\bar{x} = f_1^{\mathrm{b}}$; that is, $f_1^{\mathrm{b}} + Sf_1^{\mathrm{b}}\tau_1 = 1$, i.e. $f_1^{\mathrm{b}} = \frac{1}{1+S\tau_1}$. Finally, we know the Nash flow $f_1^{\mathrm{b}} = \frac{1}{1+S\tau_1}$ is unique since the cost on link 1 is increasing, and the cost on link 2 is constant. ■

Given this characterization of toll-influenced flows, we can now present the tolls which an infrastructure manager would charge in the absence of an arbitrageur:

Proposition 3.2: When no arbitrageur is present, the fixed and marginal-cost toll designs for link 1 which minimize total latency for the considered user population with parameter $S > 0$ are

$$\tau_1^{\mathrm{fixed}}(f_1) = \frac{1}{S}, \tag{12}$$

and

$$\tau_1^{\mathrm{mc}}(f_1) = \frac{2f_1}{S}, \tag{13}$$

respectively.

Proof: Since the congestion-optimal network flow is $f^* := (1/2, 1/2)$, the proof is obtained by substituting the tolling functions (12) and (13) into the expressions given in Proposition 3.1 and verifying that tolls satisfying (12) and (13) result in $f_1^{\mathrm{b}} = 1/2$. ■

D. Arbitrage With Fixed Tolls

Since most drivers in the network have a bias function *other than* the neutral bias function, some drivers are not behaving optimally from an objective "neutral" point of view. Since $\tau_2(f_2) = 0$, it always holds that $J_2^{\mathrm{n}}(f^*; \tau) = 1$. For arbitrary fixed tolls τ_1^{f} levied on link 1, and for $S > 0$, Proposition 3.1 admits the neutral cost of link 1 at the biased Nash flow is

$$J_1^{\mathrm{n}}(f_1^{\mathrm{b}}; \tau_1^{\mathrm{f}}) = \frac{1}{1 + S\tau_1^{\mathrm{f}}} + \tau_1^{\mathrm{f}}, \tag{14}$$

which is less than 1 (the neutral cost of link 2) when $\tau_1^{\mathrm{f}} < 1 - \frac{1}{S}$, and is greater than 1 when $\tau_1^{\mathrm{f}} > 1 - \frac{1}{S}$.

Thus, when $\tau_1^{\mathrm{f}} < 1 - \frac{1}{S}$, an arbitrageur could inform users of link 2 (i.e., high-toll-sensitivity users experiencing a neutral cost of 1) that they could experience a cost savings by switching to link 1, whose neutral cost is (14). It happens that as informed users switch to link 1, the latency of link 1 rises — so in turn, some uninformed users deviate from link 1 to link 2. Likewise, when $\tau_1^{\mathrm{f}} > 1 - \frac{1}{S}$, an arbitrageur can then inform users of link 1 (i.e., low-toll-sensitivity users experiencing a neutral cost of $\frac{1}{1+S\tau_1^{\mathrm{f}}} + \tau_1^{\mathrm{f}} > 1$) that they could experience cost savings by switching to link 2, whose neutral cost is 1. As informed users switch to link 2, the latency of link 1 decreases, leading uninformed users to deviate from link 2 to link 1.

Our next proposition demonstrates this formally: if low-enough fixed tolls τ_1^{f} are applied to link 1, the arbitrageur can recommend that at least some drivers profitably switch from link 2 to link 1, and if high-enough fixed tolls τ_1^{f} are applied to link 1, the arbitrageur can recommend that at least some drivers profitably switch from link 1 to link 2. Note that in this and following sections, we abuse notation and write $M > 0$ to denote the *mass* of drivers that the arbitrageur recommends switch from link 2 to link 1 or vice versa—which will always be made clear from context. Finally, we write $\bar{x}^{\mathrm{i}}$ to denote a threshold value of x; in informed Nash flow f^{i}, if a driver has index $x < \bar{x}^{\mathrm{i}}$, then that driver strictly prefers link 1; if a driver has index $x > \bar{x}^{\mathrm{i}}$, then that driver strictly prefers link 2.

Proposition 3.3: Suppose the network has fixed tolls τ_1^{f} levied on link 1, and let $S > 0$. Then $\Delta(M; \tau_1^{\mathrm{f}}) > 0$ if and only if M is in the interval

$$M \in \left(0, \bar{M}\right), \tag{15}$$

where $\bar{M} = \min\left\{1, \left|1 - \frac{1}{S} - \tau_1^{\mathrm{f}}\right|\right\}$. If $\tau_1^{\mathrm{f}} < 1 - \frac{1}{S}$, then $\bar{M} = 1 - \frac{1}{S} - \tau_1^{\mathrm{f}}$, so that if M satisfies (15), then the following hold:

$$f_1^{\mathrm{i}} = \frac{1 + MS\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}}, \text{ and} \tag{16}$$

$$\bar{x}^{\mathrm{i}} = \frac{1 - M}{1 + S\tau_1^{\mathrm{f}}}. \tag{17}$$

If $\tau_1^{\mathrm{f}} > 1 - \frac{1}{S}$, then $\bar{M} = 1$ or $\bar{M} = \tau_1^{\mathrm{f}} + \frac{1}{S} - 1$, so that if M satisfies (15), then the following hold:

$$f_1^{\mathrm{i}} = \frac{1 - MS\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}}, \text{ and} \tag{18}$$

$$\bar{x}^{\mathrm{i}} = \frac{1 + M}{1 + S\tau_1^{\mathrm{f}}}. \tag{19}$$

Proof: Let $S > 0$. We first assume $\tau_1^{\mathrm{f}} < 1 - \frac{1}{S}$. To obtain the proof in this case, we treat the arbitrageur's recommendation as mandatory and directly transfer M units of traffic from link 2 to link 1. Since link 1 is congestible, this transfer increases the link's cost, causing some drivers who used link 1 in f^{b} to deviate to link 2 in f^{i}. Note that the drivers who make this switch will in general be the most toll-sensitive drivers on the link, leaving $\bar{x}^{\mathrm{i}}$ as the most-sensitive driver who remains on link 1 in f^{i}. Thus, it must hold that

$$f_1^{\mathrm{i}} = M + \bar{x}^{\mathrm{i}} \tag{20}$$

drivers are selecting link 1 in f^{i}. If $\bar{x}^{\mathrm{i}} > 0$, it must be that

$$f_1^{\mathrm{i}} + S\bar{x}^{\mathrm{i}}\tau = 1 \tag{21}$$

in order to ensure that Proposition 3.1 holds. Thus, it can be verified that (16) and (17) are consistent with (20) and (21). It remains to show that it is a profitable deviation for informed drivers to switch from link 2 to link 1. That is, show that $\Delta_{21}(M, \tau_1^{\mathrm{f}}) > 0$. Now,

$$\Delta_{21}(M, \tau_1^{\mathrm{f}}) = 1 - \tau_1^{\mathrm{f}} - \frac{1 + MS\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}}. \tag{22}$$

Notice that (22) is decreasing in M. Since $\tau_1^{\mathrm{f}} < 1 - 1/S$, we have that the upper bound in (15) is $\bar{M} = 1 - \frac{1}{S} - \tau_1^{\mathrm{f}}$, so

$$\Delta_{21}(\bar{M}, \tau_1^{\mathrm{f}}) = 1 - \tau_1^{\mathrm{f}} - \frac{1 + \bar{M}S\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}} = 0.$$

Since $\Delta_{21}(M, \tau_1^{\mathrm{f}})$ is decreasing in M, and $M < \bar{M}$, we have that $0 = \Delta_{21}(\bar{M}, \tau_1^{\mathrm{f}}) < \Delta_{21}(M, \tau_1^{\mathrm{f}})$.

Next, assume $\tau_1^{\mathrm{f}} > 1 - \frac{1}{S}$. We again treat the arbitrageur's recommendation as mandatory and directly transfer M mass of traffic from link 1 to link 2. Since link 1 is congestible, the transfer *decreases* the link's cost, causing uninformed drivers who used link 2 in f^{b} to deviate to link 1 in f^{i}. The uninformed drivers that make this switch will, in general, be the least toll-sensitive drivers on the link, leaving $\bar{x}^{\mathrm{i}}$ as the least sensitive driver on link 1 in f^{i}. Thus, it holds that

$$f_1^{\mathrm{i}} = \bar{x}^{\mathrm{i}} - M \tag{23}$$

drivers are selecting link 1 in f^{i}. Once again, if $\bar{x}^{\mathrm{i}} > 0$, it must be that (21) holds so that Proposition 3.1 is ensured. Hence, it can be verified that (18) and (19) are consistent with (23) and (21). We now need to show that it is profitable for informed drivers to switch from link 1 to link 2. Hence, we need $\Delta_{12}(M, \tau_1^{\mathrm{f}}) > 0$, where

$$\Delta_{12}(M, \tau_1^{\mathrm{f}}) = \frac{1 - MS\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}} + \tau_1^{\mathrm{f}} - 1. \tag{24}$$

Notice that $\Delta_{12}(M, \tau_1^{\mathrm{f}})$ is decreasing in M. Since $\tau_1^{\mathrm{f}} > 1 - 1/S$, the upper bound in (15) is $\bar{M} = \tau_1^{\mathrm{f}} + \frac{1}{S} - 1$ or $\bar{M} = 1$. Assume first that $\bar{M} = \tau_1^{\mathrm{f}} + \frac{1}{S} - 1$, so

$$\Delta_{12}(\bar{M}, \tau_1^{\mathrm{f}}) = \frac{1 - \bar{M}S\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}} + \tau_1^{\mathrm{f}} - 1 = 0.$$

Since $\Delta_{12}(M, \tau_1^{\mathrm{f}})$ is decreasing in M, and since $M < \bar{M}$, it follows that $0 = \Delta_{12}(\bar{M}, \tau_1^{\mathrm{f}}) < \Delta_{12}(M, \tau_1^{\mathrm{f}})$. Next, assume $\bar{M} = 1$. To see that $\Delta_{12}(M, \tau_1^{\mathrm{f}}) > 0$ in this case, notice that by (24), $\Delta_{12}(M, \tau_1^{\mathrm{f}})$ is still decreasing in M. It is also the case that $\bar{M} \leq \tau_1^{\mathrm{f}} + \frac{1}{S} - 1$. Hence, the following holds:

$$0 = \Delta(M^{\times}, \tau_1^{\mathrm{f}}) \leq \Delta(\bar{M}, \tau_1^{\mathrm{f}}) < \Delta(M, \tau_1^{\mathrm{f}}), \tag{25}$$

where $M^{\times} = \tau_1^{\mathrm{f}} + \frac{1}{S} - 1$. ∎

The following proposition characterizes the total latency (1) that an infrastructure manager would be concerned with minimizing in the presence of an arbitrageur, and the revenue (6) that an arbitrageur would be concerned with maximizing.

Proposition 3.4: Suppose the network has fixed tolls τ_1^{f} levied on link 1, and let $S > 0$. Then, if $\tau_1^{\mathrm{f}} < 1 - \frac{1}{S}$, the total latency and revenue are:

$$\mathcal{L}(M, \tau_1^{\mathrm{f}}) = \frac{\left(M^2 - M + 1\right)\left(S\tau_1^{\mathrm{f}}\right)^2 + (M+1)S\tau_1^{\mathrm{f}} + 1}{\left(1 + S\tau_1^{\mathrm{f}}\right)^2} \tag{26}$$

and

$$R(M; \tau_1^{\mathrm{f}}) = \frac{M\tau_1^{\mathrm{f}}(S - MS - S\tau_1^{\mathrm{f}} - 1)}{1 + S\tau_1^{\mathrm{f}}}, \tag{27}$$

and if $\tau_1^{\mathrm{f}} > 1 - \frac{1}{S}$, the total latency and revenue are:

$$\mathcal{L}(M, \tau_1^{\mathrm{f}}) = \frac{\left(M^2 + M + 1\right)\left(S\tau_1^{\mathrm{f}}\right)^2 + (1-M)S\tau_1^{\mathrm{f}} + 1}{\left(1 + S\tau_1^{\mathrm{f}}\right)^2}, \tag{28}$$

and

$$R(M; \tau_1^{\mathrm{f}}) = \frac{M\tau_1^{\mathrm{f}}(S\tau_1^{\mathrm{f}} - MS - S + 1)}{1 + S\tau_1^{\mathrm{f}}}. \tag{29}$$

Proof: Let $S > 0$, and let f_1^{i} be an informed Nash flow. Then the total latency defined by (1) becomes

$$\mathcal{L}(M, \tau_1^{\mathrm{f}}) = {f_1^{\mathrm{i}}}^2 + \left(1 - f_1^{\mathrm{i}}\right), \tag{30}$$

and the revenue defined by (6) is

$$R(M; \tau_1^{\mathrm{f}}) = \ell_{\underline{p}}(f) + \tau_{\underline{p}}^{\mathrm{f}} - \left(\ell_{\overline{p}}(f) + \tau_{\overline{p}}^{\mathrm{f}}\right), \tag{31}$$

where $\underline{p}$ is the link informed users are leaving, and $\overline{p}$ is the link they are changing on to. Recall $\tau_2^{\mathrm{f}} = 0$. If $\tau_1^{\mathrm{f}} < 1 - 1/S$, then an arbitrageur can inform users to move from link 2 to link 1. Hence, $\underline{p} = 2$ and $\overline{p} = 1$ and

$$f_1^{\mathrm{i}} = \frac{1 + MS\tau_1^{\mathrm{f}}}{1 + S\tau_1^{\mathrm{f}}}, \tag{32}$$

by Proposition 3.3. So, (26) and (27) are obtained by substituting (32) into (30) and (31). Similarly, if $\tau_1^{\mathrm{f}} > 1 - 1/S$, then an arbitrageur can inform users to move from link 1 to link 2. Hence, $\underline{p} = 1$ and $\overline{p} = 2$ and (28) and (29) are obtained similarly. ■

E. Arbitrage with marginal-cost tolls

Now, we consider the case of arbitrage in the presence of Marginal-cost tolls. Here, to present initial results comparing the performance of marginal-cost tolls with that of fixed tolls, we focus specifically on the case that the marginal-cost toll scalar K is sufficiently low that the arbitrageur always recommends that some mass of drivers switch from link 2 to link 1.

Accordingly, suppose that the social planner levies tolls on link 1 of the form

$$\tau_1^{\mathrm{mc}}(f_1) = Kf_1, \tag{33}$$

where $K \in (0, S-1)$.

It might be hoped that these flow-varying tolls would adaptively increase the toll on link 1, acting as a deterrent to the arbitrageur. However, numerical experiments can show that this need not be the case.

The following proposition provides the form of influenced Nash flows under the influence of an arbitrageur and marginal-cost tolls.

Proposition 3.5: Let $K < S - 1$ and $\tau_1^{\mathrm{mc}}(f_1) = Kf_1$. The following hold for any $M \in [0,1]$ mass of users which the arbitrageur recommends switch from link 2 to link 1:

$$\bar{x}^{\mathrm{i}} = \frac{-1 - SKM + \sqrt{(1-SKM)^2 + 4SK}}{2SK}, \text{ and} \tag{34}$$

$$f_1^{\mathrm{i}} = M + \frac{-1 - SKM + \sqrt{(1-SKM)^2 + 4SK}}{2SK}. \tag{35}$$

Proof: A Nash flow for the neutral population has $f_1^{\mathrm{n}} = 1/(1+K)$. It can be verified from Proposition 3.1 that when $K < S-1$, the arbitrage-free Nash flow has $f_1^{\mathrm{b}} < 1/(1+K)$, and thus that $c_1^{\mathrm{n}}(\cdot) < c_2^{\mathrm{n}}(\cdot)$. That is, when $K < S-1$, the arbitrageur can profitably recommend that some mass M of drivers deviate from link 2 to link 1.

Accordingly, let the arbitrageur recommend that mass $M \in [0,1]$ drivers deviate from link 2 to link 1. The resulting flow (35) can be computed by solving the following equation for $\bar{x}^{\mathrm{i}}$:

$$M + \bar{x}^{\mathrm{i}} + S\bar{x}^{\mathrm{i}}K(M + \bar{x}^{\mathrm{i}}) = 1. \tag{36}$$

If $M + \bar{x}^{\mathrm{i}}$ drivers use link 1, then the most-sensitive of them (with tax sensitivity $S\bar{x}^{\mathrm{i}}$) is precisely indifferent between links 1 and 2. Any user $x < \bar{x}^{\mathrm{i}}$ strictly prefers link 1, and any user $x > \bar{x}^{\mathrm{i}}$ strictly prefers link 2. Provided that $M + \bar{x}^{\mathrm{i}} \leq 1$, it holds that $f_1^{\mathrm{i}} = M + \bar{x}^{\mathrm{i}}$ is a Nash flow.

To show this, note that (36) describes a convex quadratic equation $P(x) = 0$ with at most one positive root $\bar{x}^{\mathrm{i}}$ given by (34). If $P(1-M) > 0$, then because P is convex, it must hold that $\bar{x}^{\mathrm{i}} \leq 1 - M$. This is easily verified:

$$\begin{aligned} P(x) &= SKx^2 + (1 + SKM)x - (1-M) \\ &= \Big|_{x=1-M} SK(1-M)^2 + (1+SKM)(1-M) - (1-M) \\ &= (1-M)(SK(1-M) + SKM) \geq 0, \end{aligned}$$

since $M \leq 1$. Thus, $M + \bar{x}^{\mathrm{i}} \leq 1$ for any admissible M, obtaining the proof. ■

The expressions for arbitrageur revenue and total latency are thus possible to compute directly from Proposition 3.5, but we believe the closed-form expressions offer little insight and for reasons of space we do not reproduce them here.

IV. Conclusions

This paper has formally posed a model of 3rd-party informational arbitrage in transportation networks under the influence of monetary tolls, and analyzed an instance of the problem on a simple transportation network. Our work suggests research opportunities in several directions, including analysis for more complex networks, more complex forms of tolling function, and the deterrence effects of large tolls.

References

[1] L. J. Ratliff, R. Dong, S. Sekar, and T. Fiez, "A Perspective on Incentive Design: Challenges and Opportunities," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, no. 1, pp. 1–34, 2018.

[2] T. Chen, W. Liu, Q. Fang, J. Guo, and D.-Z. Du, "Minimizing Misinformation Profit in Social Networks," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1206–1218, dec 2019.

[3] A. Tversky and D. Khaneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 1974.

[4] B. L. Ferguson, P. N. Brown, and J. R. Marden, "The Effectiveness of Subsidies and Tolls in Congestion Games," *IEEE Transactions on Automatic Control*, vol. 67, no. 6, pp. 2729–2742, jun 2022.

[5] L. Cianfanelli, G. Como, A. E. Ozdaglar, and F. Parise, "Optimal intervention in transportation networks," *IEEE Transactions on Automatic Control*, vol. 68, no. 12, pp. 7073–7088, 2023.

[6] Y. Chen and V. Gupta, "Toll Design for Routing Games With Stochastic Demands," *IEEE L-CSS*, vol. 6, pp. 3445–3450, 2022.

[7] P. Kleer and G. Schäfer, "The Impact of Worst-Case Deviations in Non-Atomic Network Routing Games," *Theory of Computing Systems*, vol. 63, no. 1, pp. 54–89, 2019.

[8] R. Cole, T. Lianeas, and E. Nikolova, "When does diversity of agent preferences improve outcomes in selfish routing?" in *IJCAI*, 2018, pp. 173–179.

[9] M. Abdallah, T. Cason, S. Bagchi, and S. Sundaram, "The Effect of Behavioral Probability Weighting in a Simultaneous Multi-Target Attacker-Defender Game," in *2021 ECC*. IEEE, 2021, pp. 933–938.

[10] L. Fleischer, K. Jain, and M. Mahdian, "Tolls for Heterogeneous Selfish Users in Multicommodity Networks and Generalized Congestion Games," in *Proc. 45th IEEE FOCS*, Rome, Italy, 2004, pp. 277–285.

[11] R. Cole, Y. Dodis, and T. Roughgarden, "Pricing network edges for heterogeneous selfish users," in *Proc. of the 35th ACM STOC*, 2003, pp. 521–530.

[12] R. D. Connors and A. Sumalee, "A network equilibrium model with travellers' perception of stochastic travel times," *Transportation Research Part B: Methodological*, vol. 43, no. 6, pp. 614–624, 2009.

[13] Y. Guan, A. M. Annaswamy, and H. Eric Tseng, "Cumulative Prospect Theory Based Dynamic Pricing for Shared Mobility on Demand Services," in *2019 IEEE 58th CDC*. IEEE, dec 2019, pp. 2239–2244.

[14] A. Mas-Colell, "On a Theorem of Schmeidler," *Mathematical Economics*, vol. 13, pp. 201–206, 1984.

[15] T. Roughgarden, *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.

[16] P. N. Brown and J. R. Marden, "Optimal mechanisms for robust coordination in congestion games," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2437–2448, 2018.