Avoiding Perverse Incentives in Affine Congestion Games

Philip N. Brown and Jason R. Marden

Abstract—In engineered systems whose performance depends on user behavior, it is often desirable to influence behavior in an effort to achieve performance objectives. However, doing so naively can have unintended consequences; in the worst cases, a poorly-designed behavior-influencing mechanism can create a perverse incentive which encourages adverse user behavior. For example, in transportation networks, marginal-cost tolls have been studied as a means to incentivize low-congestion network routing, but have typically been analyzed under the assumption that all network users value their time equally. If this assumption is relaxed, marginal-cost tolls can create perverse incentives which increase network congestion above un-tolled levels. In this paper, we prove that if some network users are unresponsive to tolls, any taxation mechanism that does not depend on network structure can create perverse incentives. Thus, to systematically avoid perverse incentives, a taxation mechanism must be network-aware to some extent. On the other hand, we show that a small amount of additional information can mitigate this negative result; for example, we show that it is relatively easy to avoid perverse incentives on parallel-path networks, and we derive the taxation mechanism that minimizes congestion for worst-case user populations.

I. INTRODUCTION

Many of today’s engineered systems are richly connected with their users; economic, social, and technical objectives are often interconnected in complex ways. Examples of this can be found in ridesharing systems [1], transportation networks [2], and power grids [3]. As the interconnections between social and engineered systems increase, the engineer’s task increasingly includes influencing the behavior of system users. Accordingly, recent research has focused on developing new analytical tools for influencing social behavior to achieve engineering objectives [4]–[8]. One intrinsic challenge in this setting is that of uncertainty in the social systems to be influenced. It can be difficult to characterize user preferences and decision-making processes in a social system, and any behavior-influencing mechanism must take this uncertainty into account. This has led to recent efforts to develop behavior-influencing mechanisms that are robust to a wide range of possible mischaracterizations [9]–[11].

One particular area of interest is that of influencing drivers’ routing choices in transportation networks. It is well-known that if individual drivers choose their own routes through a congestible network to minimize their personal delay, the resulting aggregate network delay can be significantly worse than optimal [12]. Recent research has suggested levying road taxes which modify agents’ costs and incentivize more-efficient network flows. Many such taxation schemes require the tax-designer to have a perfect characterization of all underlying system variables: network topology, road congestion characteristics, user tax-sensitivities, and overall demand. Given such a perfect characterization, it is known that a system planner can design taxes which incentivize optimal network flows [13]–[15]. Unfortunately, recent results have demonstrated that taxes designed for one problem instance can incentivize inefficient behavior on different (yet closely-related) instances, indicating that these taxes lack robustness to system mischaracterizations.

In [16], the authors propose a semantic framework for robustness as applied to behavior-influencing mechanisms. In this framework, taxes are designed for some nominal routing problem, and the problem is then perturbed in a variety of ways. For each perturbed routing problem, the congestion induced by the original taxes is then compared to two benchmarks: the optimal aggregate delay on the perturbed problem, and the delay of an un-influenced flow on the perturbed problem. The taxes are said to be strongly robust to that type of perturbation if the nominal taxes induce optimal flows on all perturbed networks. The taxes are said to be weakly robust to the perturbation if on the perturbed networks the flows incentivized by the nominal taxes are never worse than un-influenced flows. That is, if a taxation mechanism is weakly robust, the system planner can be certain that taxing is never worse than not taxing; alternatively, these taxes will never create perverse incentives.

A prominent example of a strongly-robust taxation mechanism is that of marginal-cost tolls, which are known to incentivize optimal network flows without requiring a priori knowledge of user demand or network topology provided that all users trade off time and money equally [17], [18]. An attractive feature of marginal-cost tolls is that the toll on each network link depends only on that link’s flow and congestion properties, and can be computed without any information about overall network topology. This property is known as network-agnosticity, and is a desirable characteristic of any taxation mechanism since by construction it confers robustness to variations in network structure.

However, recent research has suggested that this network-agnosticity comes at a price; if network users have unknown price-sensitivities, marginal-cost tolls fail to be strongly-robust to mischaracterizations of user price-sensitivity [16]. Unfortunately, in the most general networks, if users have diverse price-sensitivities, off-the-shelf marginal-cost tolls are not even weakly robust [16]; this is demonstrated in...
Example 2.1 in this paper. Put plainly, marginal-cost tolls can create perverse incentives if applied naively. Despite this fact, [9], [19] show that for parallel networks subject to a particular utilization constraint, scaled marginal-cost tolls (and affine tolls, a closely-related variant) are weakly robust. This indicates that if the toll-designer has additional information about the routing problem setting (e.g., information about the allowable class of networks), the weak-robustness of marginal-cost tolls may be recovered.

Accordingly, the central goal of this paper is to understand the relationship between robustness and network-agnosticity more fully, and determine specifically under what conditions network-agnostic tolls can be weakly robust. To that end, in our first and most general result we demonstrate that if some users are unresponsive to tolls, the only weakly-robust network-agnostic taxation mechanism is essentially the taxation mechanism which charges zero tolls. That is, to avoid perverse incentives, taxes must depend on some information regarding network structure.

Fortunately, a taxation mechanism need not depend on much additional information: our second result states that for the class of affine-cost parallel-path networks, a weakly-robust network-agnostic taxation mechanism always exists; that is, knowledge of the class of allowable networks can render weak robustness possible. We give a full characterization of the space of weakly-robust network-agnostic taxation mechanisms for this setting, and we derive the taxation mechanism that minimizes network congestion in worst case over any unknown user toll-sensitivities. This is encouraging from the standpoint of a toll-designer, since it suggests that a little additional information about the problem setting can greatly expand the designer’s toolbox.

II. MODEL AND RELATED WORK
A. Routing Game

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 $E \subseteq (V \times V)$. We call a source/destination vertex pair $(s_c, t_c) \in (V \times V)$ a commodity, and the set of all commodities $C$. We assume that for each $c \in C$, there is a mass of traffic $r_c > 0$ that needs to be routed from $s_c$ to $t_c$. We write $\mathcal{P}_c \subset 2^E$ to denote the set of paths available to traffic in commodity $c$, where each path $p \in \mathcal{P}_c$ consists of a set of edges connecting $s_c$ to $t_c$. Let $\mathcal{P} = \bigcup \{ \mathcal{P}_c \}$. A network is called a parallel-path network if all paths are disjoint; i.e., for all paths $p, p' \in \mathcal{P}$, $p \cap p' = \emptyset$.

A feasible flow $f \in \mathbb{R}^{|\mathcal{P}|}$ is an assignment of traffic to various paths such that for each commodity, $\sum_{p \in \mathcal{P}_c} f_p = r_c$, where $f_p \geq 0$ denotes the mass of traffic on path $p$. Without loss of generality, we assume that $\sum_{c \in C} r_c = 1$.

Given a flow $f$, the flow on edge $e$ is given by $f_e = \sum_{p \in \mathcal{P}_c} f_p$. To characterize transit delay as a function of traffic flow, each edge $e \in E$ is associated with a specific affine latency function $\ell_e : [0, 1] \rightarrow [0, \infty)$ of the form $\ell_e(f_e) = a_e f_e + b_e$, where $a_e \geq 0$ and $b_e \geq 0$ are edge-specific constants. We measure the cost of a flow $f$ by the total latency, given by

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

where $\ell_p(f_p) = \sum_{e \in p} \ell_e(f_e)$ denotes the latency on path $p$. We denote the flow that minimizes the total latency by

$$f^* \in \operatorname{argmin}_{f \text{ is feasible}} \mathcal{L}(f).$$

Due to the convexity of $\ell_e$, $\mathcal{L}(f^*)$ is unique.

A routing problem is given by $G = (V, E, C, \{ \ell_e \})$. The set of all routing problems is written $\mathcal{G}$.

To study the effect of taxes on self-interested behavior, we model the above routing problem as a non-atomic congestion game. We assign each edge $e \in E$ a flow-dependent taxation function $\tau_e : \mathbb{R}^+ \rightarrow \mathbb{R}^+$. To characterize users’ taxation sensitivities, let each user $x \in [0, r_c]$ have a taxation sensitivity $s_x^e \in [S_L, S_U] \subseteq \mathbb{R}^+$, where $S_L \geq 0$ and $S_U \leq +\infty$ are lower and upper sensitivity bounds, respectively. Note that we allow $S_U$ to take the value $+\infty$. Given a flow $f$, the cost that user $x$ experiences for using path $p \in \mathcal{P}_c$ is of the form

$$J_x(f) = \sum_{e \in p} [\ell_e(f_e) + s_x^e \tau_e(f_e)],$$

and we assume that each user selects the lowest-cost path from the available source-destination paths. We call a flow $f$ a Nash flow if for all commodities $c \in C$ and all users $x \in [0, r_c]$ we have

$$J_x(f) = \min_{p \in \mathcal{P}_c} \sum_{e \in p} [\ell_e(f_e) + s_x^e \tau_e(f_e)].$$

It is well-known that a Nash flow exists for any non-atomic congestion game of the above form [20].

We assume that the sensitivity distribution function $s$ is unknown; for a given routing problem $G$, we define the set of possible sensitivity distributions as the set of Lebesgue-measurable functions $S = \{ s^e : [0, r_c] \rightarrow [S_L, S_U] \}_{e \in C}$.

B. Taxation Mechanisms and Robustness

In this paper, we consider a particular type of taxation mechanism which we term network-agnostic. Here, each edge’s taxation function is computed using only locally-available information. That is, $\tau_e(f_e)$ depends only on $\ell_e$, not on edge $e$’s location in the network, the overall network topology, the overall traffic rate, or the congestion properties of any other edge. A network-agnostic taxation mechanism $\tau$ is thus a mapping from latency functions to taxation functions, and any specific edge taxation functions is given by

$$\tau_e(\cdot) = \tau(\ell_e).$$

To evaluate taxation mechanisms, we adopt the robustness framework introduced in [16] in which a distinction is drawn between strong and weak robustness. In either case, for each network we compare the worst-case Nash congestion induced by a particular taxation mechanism to two benchmarks: first, the optimal total latency on the network; second, the total latency of an un-influenced Nash flow on the network. Formally, in the context of network-agnostic taxation mechanisms, we write $\mathcal{L}^\text{nf}(G, s, \tau)$ to denote the total latency of
a Nash flow for routing problem \( G \) and population \( s \) induced by taxation mechanism \( \tau \). We write \( \mathcal{L}^{uf}(G, \emptyset) \) to denote the total latency of an un-influenced Nash flow (note that when there are no tolls, the sensitivity distribution plays no role), and \( \mathcal{L}^*(G) = \mathcal{L}(f^*) \) to denote the optimal latency.

Taxation mechanism \( \tau \) is said to be strongly robust if for every network it induces optimal Nash flows for any sensitivity distribution; that is, for all \( G \in \mathcal{G} \),

\[
\sup_{s \in S} \mathcal{L}^{uf}(G, s, \tau) = \mathcal{L}^*(G). \tag{6}
\]

On the other hand, \( \tau \) is said to be weakly robust if for every network and sensitivity distribution, the total latency induced by \( \tau \) never exceeds the total latency of an un-influenced Nash flow; i.e., for all \( G \in \mathcal{G} \),

\[
\sup_{s \in S} \mathcal{L}^{uf}(G, s, \tau) \leq \mathcal{L}^{uf}(G, \emptyset). \tag{7}
\]

Weak robustness is a guarantee that perverse incentives will never arise; note that at a minimum, it is always trivially true that the zero-toll is weakly robust.

C. Related Work and Examples

The following is a brief survey of relevant work on the robustness of taxation mechanisms in congestion games. One type of taxation mechanism which generally fails to be robust is that of fixed tolls, which for any \( e \in G \), \( \tau_e(f_e) = q_e \) for some \( q_e \geq 0 \). If network, traffic-rate, and user sensitivity specifications are known precisely, it is possible to compute fixed tolls which induce optimal Nash flows on any network [13], [14]. However, if any of these pieces of information are unknown, fixed tolls fail to be strongly robust; if fixed tolls are further restricted to be network-agnostic, they fail even to be weakly-robust [16]. One variant is dynamic fixed tolls, which can be adjusted periodically by a central authority in response to user behavior. In [21] a dynamic fixed tolling scheme is proposed which can incentivize desired flows for homogeneous populations, even if the network’s latency functions are unknown. The central computation means these tolls are not network-agnostic.

A classic example of a strongly-robust network-agnostic taxation mechanism is that of the marginal-cost or Pigovian taxation mechanism \( \tau^{mc} \). For any \( e \in G \) with latency function \( \ell_e \), the accompanying marginal-cost toll is

\[
\tau^{mc}_e(f_e) = f_e \cdot \ell'_e(f_e), \quad \forall f_e \geq 0, \tag{8}
\]

where \( \ell' \) represents the flow derivative of \( \ell \). In [17] the authors show that for any \( G \in \mathcal{G} \), it is true that \( \mathcal{L}^*(G) = \mathcal{L}^{uf}(G, \emptyset, \tau^{mc}) \), provided that all users have a sensitivity equal to 1. Thus, by construction, marginal-cost tolls are strongly-robust to perturbations of network structure and traffic rate, since each taxation function has no dependence on either. Unfortunately, this robustness depends strongly on the homogeneity of the user population. The following example demonstrates that marginal-cost tolls are not even weakly robust in multicommodity networks for heterogeneous populations.

Example 2.1: Consider the network depicted in Figure 1. There are two source nodes; 0.5 units of traffic from the upper source with sensitivity \( s_1 \) share a common destination with 1 unit of traffic from the lower source with sensitivity \( s_2 \). It is simple to verify that if all traffic trades off time and money equally (i.e., \( s_1 = s_2 = 1 \)), marginal-cost tolls incentivize the optimal flow depicted on the left of the figure. However, if the upper-source traffic keeps \( s_1 = 1 \) but the lower-source traffic has \( s_2 = 0 \) (i.e., they care only about time), marginal-cost tolls create a perverse incentive, resulting in the configuration depicted on the right – which exhibits higher total latency than the un-tolled configuration. This proves the lack of weak robustness of marginal-cost tolls in multicommodity heterogeneous networks.

Moving past marginal-cost tolls, the authors of [19] exhibit a pair of network-agnostic taxation mechanisms which are proven to be weakly-robust under particular conditions. The first taxation mechanism, called the universal taxation mechanism \( \tau^u \), assigns tolling functions of

\[
\tau^u_e(f_e) = \kappa_u(f_e, b_e, f_e) \tag{9}
\]

and the authors show that for any network, if \( S_L > 0 \), these universal tolls are weakly-robust for large-enough \( \kappa_u \). Here, \( S_L > 0 \) implies that user sensitivities are bounded away from zero; the \( \kappa_u \) required for weak robustness depends on \( S_L \) in general. However, if the tax designer does not know a lower bound on user sensitivities (i.e., \( S_L = 0 \), it is possible to show on the network in Example 2.1 that \( \tau^u \) fails to be weakly robust.

The second taxation mechanism of [19] demonstrates the principle that additional information about the class

\footnotetext{1}{Here, we present only the universal tolls associated with affine latency functions; [19] exhibits a generalization which applies to any convex latency function.}
of problems can be exploited to influence behavior more effectively while preserving weak robustness. For the class of parallel networks, if $S_L$ and $S_U$ are known and positive, and the un-tolled Nash flow of every network has positive flow on every link (this latter condition is termed fully-utilized), tolls of the following form are weakly-robust for any $\kappa_A \geq 0$:

$$\tau^A(f_c) = \kappa_A a_c f_c + \max \left\{ 0, \frac{\kappa_A S_L S_U - 1}{S_L + S_U + 2\kappa_A S_L S_U} b_e \right\}.$$  

(10)

However, if $\tau^A$ is applied to a network that is not fully-utilized, it can easily create perverse incentives, as we show in the following example.

**Example 2.2:** Figure 2 depicts the canonical network known as Pigou’s Example, a two-link parallel network with $\ell_1(f_1) = f_1$ and $\ell_2(f_2) = 1$. Suppose $S_L = 0.1$ and $S_U = 10$; we will charge tolls to this network according to $\tau^u$ and $\tau^A$ (with $\kappa_u = \kappa_A = 2$) and record the worst-case total latency for various traffic rates.

For universal tolls, it can be shown that for any $r > 0$, the worst-case Nash flows on this network occur for a homogeneous population with $s \equiv S_L$. The opposite is true for $\tau^A$; worst-case Nash flows occur for a homogeneous population with $s \equiv S_U$. Using these facts, for each taxation mechanism $\tau$, the quantity $\sup_{s \in [S_L, S_U]} \mathcal{L}^{\text{nil}}(G, s, \tau)$ is plotted as a function of $r$ in Figure 2. On the plot, a value greater than 1 indicates a perverse incentive. Note that for high traffic (high values of $r$), $\tau^A$ outperforms $\tau^u$. However, the good performance guarantees of $\tau^A$ come at a price: for low values of $r$ (i.e., the network is not fully-utilized), $\tau^A$ induces Nash flows that have nearly twice the total latency of an un-tolled flow. This poignantly illustrates that the full-utilization assumption employed by $\tau^A$ is not without loss of generality.

**III. OUR CONTRIBUTIONS**

**A. Impossibility in Multicommodity Networks**

Our first result considers the case that a toll-designer wishes to design a network-agnostic taxation mechanism that improves the efficiency of Nash flows on any network. Theorem 3.1 shows that this is impossible on multicommodity networks if user sensitivities are heterogeneous and are not bounded away from 0. That is, any non-trivial network-agnostic taxation mechanism can create perverse incentives.

**Theorem 3.1:** For the class of all multicommodity affine-latency networks, if $S_L = 0$ and $S_U > 0$, a network-agnostic taxation mechanism $\tau$ is weakly robust if and only if for every network $G \in \mathcal{G}$ and population $s$,

$$\mathcal{L}^{\text{nil}}(G, s, \tau) = \mathcal{L}^{\text{nil}}(G, \emptyset).$$

(11)

Alternatively, if a network-agnostic taxation mechanism improves congestion for one routing problem, it must degrade congestion for another. As a first step towards proving

---

2Note that a taxation mechanism that possesses this property always assigns taxes of $\tau_c(f_c) = \mu \mu(f_c)$ for some $\mu \geq 0$; tolls of this form have no influence on any Nash flow.
Theorem 3.1, in Lemma 3.2 we present necessary conditions for the weak robustness of any network-agnostic taxation mechanism; we will subsequently show that any taxation mechanism satisfying these conditions will fail to be weakly robust in certain multicommodity routing problems.

Lemma 3.2: For networks with affine cost functions, if a network-agnostic taxation mechanism is weakly robust, it assigns taxes given by

$$\tau_e(f_e) = \kappa_1 a_e f_e + \kappa_2 b_e,$$

(12)

for some $$\kappa_1 \geq 0$$ and $$\kappa_2 \geq 0$$, and for every possible user $$x \in [0, r]$$ coefficients $$\kappa_1$$ and $$\kappa_2$$ satisfy

$$\frac{s_x (\kappa_1 - \kappa_2)}{1 + s_x \kappa_2} \in [0, 1].$$

(13)

The proof of Lemma 3.2 appears in the Appendix.

1) Proof of Theorem 3.1: We prove this by showing on the network of Example 2.1 that if users are heterogeneous and price-sensitivities can take a value of 0, no tolls satisfying the conditions of Lemma 3.2 can be weakly-robust. Consider the network depicted in Figure 1. A high-sensitivity population with sensitivity $$s_1$$ and mass 1/2 is traveling from the upper source; a low-sensitivity population with sensitivity $$s_2$$ and mass 1 is traveling from the lower source; the populations share a common destination.

Charge tolls on this network in accordance with Lemma 3.2 (that is, specify $$\kappa_1$$ and $$\kappa_2$$ satisfying (13) for both $$s_x = s_1$$ and $$s_x = s_2$$), and define $$\gamma_1$$ and $$\gamma_2$$ as follows:

$$\gamma_1 = \frac{s_1 (\kappa_1 - \kappa_2)}{1 + s_1 \kappa_2},$$

and

$$\gamma_2 = \frac{s_2 (\kappa_1 - \kappa_2)}{1 + s_2 \kappa_2}.$$ (14)

The upper population’s Nash-flow incentive constraint induced by these tolls is given by

$$(1 + \gamma_1) f_1 + 1 = (1 + \gamma_1) f_2,$$ (15)

and the lower population’s by

$$(1 + \gamma_2) f_2 = 1.$$ (16)

It can be shown that any Nash flow for which $$\gamma_2 \leq \gamma_1$$ must satisfy (15) and (16); since the system of equations is upper-triangular, $$f_2$$ depends only on $$\gamma_2$$:

$$f_1 = \frac{1}{1 + \gamma_2} - \frac{1}{1 + \gamma_1},$$

and

$$f_2 = \frac{1}{1 + \gamma_2}.\ (17)$$

In essence, the low-sensitivity population holds all the power; the flow on path 2 is not affected by the sensitivities of the high-sensitivity population.

Consider the definitions of $$\gamma_1$$ and $$\gamma_2$$; note that for any fixed choice of $$\kappa_1$$ and $$\kappa_2$$, $$\gamma_2$$ can be made arbitrarily-close to 0 by choosing a very low $$s_2$$. To model the extreme case, let $$s_2 = 0$$ so that $$\gamma_2 = 0$$. Then $$f_2 = 1$$ and the total latency on the network as a function of $$\gamma_1 > 0$$ is given by

$$L(\gamma_1) = 1 + \frac{\gamma_1^2}{1 + \gamma_1} + \frac{\gamma_1}{1 + \gamma_1} + \frac{1 - \gamma_1}{2(1 + \gamma_1)}.$$

$$= \frac{5\gamma_1^2 + 6\gamma_1 + 4}{2(1 + \gamma_1)^2} > 1.5.$$

Thus, charging tolls that induce $$\gamma_1 > 0$$ cause the total latency of a tolled Nash flow to be greater than that of the un-tolled Nash flow. The only tolls that guarantee $$\gamma_1 = 0$$ have tolling coefficients $$\kappa_1 = \kappa_2$$; any tolls of this form have $$L^u(G, s, \tau) = L^u(G, \emptyset)$$. ■

B. Weakly Robust Tolls for Parallel-Path Networks

We now ask if knowledge of a small amount of information can mitigate the negative result of the previous section. Indeed, if networks are known to be single-commodity parallel-path networks, Theorem 3.3 shows that weakly-robust network-agnostic taxation mechanisms always exist. This is true even in extreme cases when $$S_L = 0$$ or $$S_U = \infty$$. The implications for a toll-designer are encouraging: this result demonstrates the intuitive principle that information regarding the possible class of networks can greatly expand the designer’s toolbox.

Before stating the theorem, we point out that worst-case performance guarantees provided by a taxation mechanism can often be improved by increasing all edge tolls appropriately (as discussed in [19]). Thus, in order to make meaningful statements about congestion-minimizing tolls, it is useful to parameterize tolls by a stylized upper-bound; the parameter $$\kappa_{\text{max}} > 0$$ plays this role in the following theorem.

Theorem 3.3: For single-commodity parallel-path networks with affine latency functions, a network-agnostic taxation mechanism is weakly robust if and only if it satisfies the conditions of Lemma 3.2. Furthermore, for any $$S_U < \infty$$ and toll-scalar upper-bound $$\kappa_{\text{max}} \geq 0$$ the congestion-minimizing taxation mechanism assigns tolling functions

$$\tau_e(f_e) = \kappa_{\text{max}} a_e f_e + b_e \max \left\{0, \frac{\kappa_{\text{max}} S_U - 1}{2S_U}\right\}.$$ (18)

If $$S_U = +\infty$$, then for any $$S_L \geq 0$$, (18) simplifies to

$$\tau_e(f_e) = \kappa_{\text{max}} \left(a_e f_e + \frac{b_e}{2}\right).$$ (19)

The proof of Theorem 3.3 appears in the Appendix.

Here, we find that the universal tolls of [19] are in fact weakly robust for parallel-path networks. Another important fact to note is that (18) gives some insight into the robustness of scaled marginal-cost tolls: Recall that for affine-latency congestion games, marginal-cost tolls are given by $$\tau_{e\text{mc}}(f_e) = a_e f_e$$, and incentivize optimal Nash flows for unit-sensitivity homogeneous populations. If $$\kappa_{\text{max}} = 1/S_U$$, (18) gives the congestion-minimizing scaled marginal-cost toll as $$\tau_e(f_e) = a_e f_e / S_U$$. This can be interpreted as a conservatively-scaled marginal-cost toll; it implies that the best way to avoid perverse incentives with marginal-cost tolls is to charge tolls as though all users have sensitivity equal to $$S_U$$.

IV. CONCLUSION

In this paper, we have presented initial findings on the weak robustness of network-agnostic taxation mechanisms; we showed that in general routing problems, network-agnosticity carries the risk of perverse incentives if some
network users are unresponsive to tolls. On the other hand, we showed that on parallel-path networks, perverse-incentives can be systematically avoided, and we characterized the full space of weakly robust network-agnostic taxation mechanisms for this setting. Throughout, we gauged everything from a draconian worst-case perspective; relaxing this approach slightly may yield significant robustness gains. For example, if a tax-designer has coarse distributional knowledge of a user population’s price sensitivities, it is possible that this knowledge can be exploited to employ more-aggressive taxation schemes which remain weakly-robust. Characterizing these tradeoffs between information and performance is the subject of ongoing research.

REFERENCES


APPENDIX: PROOFS

A. Proof of Lemma 3.2

We shall prove Lemma 3.2 with a series of simple example networks; in each case, the approach will be to take a taxation mechanism that fails to satisfy one of the conditions, and construct a network on which it creates perverse incentives.

Proof of Lemma 3.2: First, we show that for networks with affine cost functions, any weakly-robust network-agnostic taxation mechanism assigns taxes of the simple affine form

$$\tau_e(f_e) = \kappa_1 a_e f_e + \kappa_2 b_e.$$  \hspace{1cm} (20)

A network-agnostic taxation mechanism is fundamentally a mapping from latency functions to taxation functions. Let $$\tau^a$$ be a weakly-robust taxation mechanism; that is, for any affine latency function $$\ell_e(f_e) = a_e f_e + b_e$$, the taxation function on edge $$e$$ is given by $$\tau_e(f_e) = \tau^a(a_e f_e + b_e)$$. Consider the network in Figure 3(a): let $$\ell_1(f_1) = a_1 f_1$$, $$\ell_2(f_2) = b_2 f_2$$, and $$\ell_3(f_3) = a_3 f_3 + b_3$$, for any choice of $$a$$ $\geq 0$$ and $$b \geq 0$$, the un-tolled Nash flow on this network splits the traffic evenly between the upper path and the lower path; this flow is also optimal. Thus, weakly-robust tolls must charge the same total amount on the upper path as they do on the lower path, or

$$\tau^a(a f + b) = \tau^a(a f) + \tau^a(b).$$ \hspace{1cm} (21)

Repeat this exercise with $$\ell_1(f_1) = a_1 f_1$$ and $$\ell_2(f_2) = a_2 f_2$$ to see that

$$\tau^a(a_1 f_1 + a_2 f_2) = \tau^a(a_1 f_1) + \tau^a(a_2 f_2).$$ \hspace{1cm} (22)

Lastly, the same procedure with all constant latency functions $$\ell_1(f_1) = b_1$$ and $$\ell_2(f_2) = b_2$$ shows that

$$\tau^a(b_1 + b_2) = \tau^a(b_1) + \tau^a(b_2).$$ \hspace{1cm} (23)

Since $$\tau^a$$ is a continuous mapping, the additivity of (21), (22), and (23) further imply that $$\tau^a$$ is linear, or there exist scalars $$\kappa_1 \in \mathbb{R}^+$$ and $$\kappa_2 \in \mathbb{R}^+$$ such that

$$\tau^a(a f + b) = \kappa_1 a f + \kappa_2 b.$$ \hspace{1cm} (24)

Next, we can use this additivity property to show (13). For a user $$x$$ with sensitivity $$s_x$$, tolls as defined in (20) induce edge cost functions

$$J_x^e(f_e) = a_e f_e + b_e + s_x (\kappa_1 a_e f_e + \kappa_2 b_e),$$

$$= (1 + s_x \kappa_2) a_e f_e + b_e + s_x (\kappa_1 - \kappa_2) a_e f_e.$$ \hspace{1cm} (25)

Since dividing all cost functions by a finite (even user-specific) scalar does not change the associated Nash flows, we can divide all cost functions by $$(1 + s_x \kappa_2)$$ to obtain the equivalent cost functions

$$J_x^e(f_e) = a_e f_e + b_e + s_x (\kappa_1 - \kappa_2) a_e f_e \frac{1 + s_x \kappa_2}{1 + s_x \kappa_2}.$$ \hspace{1cm} (26)

For any $$\kappa_1$$, $$\kappa_2$$, and population $$s$$ the following user-specific coefficient is always well-defined as a real number:

$$\gamma_x = \frac{s_x (\kappa_1 - \kappa_2)}{1 + s_x \kappa_2},$$ \hspace{1cm} (27)
However, if this network with sufficient to increase the total latency; showing that tolls with 

$$\gamma > x \in \mathbb{R}.$$  

In the following analysis, we assume that the population is homogeneous, so that for all \(x, \gamma_x = \gamma\).

First, suppose that \(\gamma < 0\) (i.e., \(\kappa_1 < \kappa_2\)). We will show this is a perverse incentive in any network of the form shown in Figure 3(b). This network has two links in parallel; link 1 has linear latency function \(l_1(f_1) = a f_1\), link 2 has constant latency function \(l_2(f_2) = b\), where \(a > 0\) and \(b > 0\), and \(r > b/a\) so that the un-tolled Nash flow \(f^{\text{nf}}\) is given by

$$f_1^{\text{nf}} = \frac{b}{a}, \quad \text{and} \quad f_2^{\text{nf}} = r - \frac{b}{a},$$

with a total latency of \(L(f^{\text{nf}}) = br\). However, \(\gamma_x < 0\) decreases the cost of edge 1, inducing a tolled Nash flow with \(f_1^{\text{toll}} > f_1^{\text{nf}}\). Any increase in \(f_1\) is sufficient to increase the total latency; showing that tolls with \(\gamma < 0\) cannot be weakly robust.

Second, suppose that \(\gamma > 1\). Consider again the network in Figure 3(b); this time, let \(a = b = 1\). This is the classical Pigou’s Network, typically used to demonstrate the inefficiency of uninfluenced Nash flows. Here, we use it to demonstrate the inefficiency of over-influenced Nash flows. Consider this network with \(r = 1/2\); at this traffic rate, the un-tolled Nash flow and the optimal flows are equal:

$$f^{\text{nf}} = f^{\text{opt}} = (1/2, 0) \quad \text{with total latency} \quad L(f^{\text{opt}}) = 1/4.$$

However, if \(\gamma > 1\), the resulting tolled Nash flow has

$$f^{\text{toll}} = \left(\frac{1}{1+\gamma}, \frac{\gamma-1}{2(1+\gamma)}\right).$$

When \(\gamma > 1\), the total latency of this flow is given by

$$L(f^{\text{toll}}) = \left(\frac{1}{1+\gamma}\right)^2 + \frac{\gamma - 1}{2(1+\gamma)} \geq 1/4.$$

Thus, affine tolls can only be weakly-robust if for all \(x, \gamma_x \in [0, 1]\), completing the proof.

**B. Proof of Theorem 3.3 and associated Lemmas**

To facilitate our arguments, we assign labels of \(i \in \{1, \ldots, |P|\}\) to a network’s paths such that if \(b_i = \sum_{e \in P_i} b_e\), for all \(i\) we have \(b_i < b_{i+1}\). Similarly, we write \(f_i = \sum_{e \in P_i} f_e\) and \(a_i = \sum_{e \in P_i} a_e\). We define the matrix \(A \in \mathbb{R}^{|P||P|}\) as the diagonal matrix with all \(a_i\) coefficients on the diagonal, and the column vector \(b \in \mathbb{R}^{|P|}\) as the vector of \(b_i\) coefficients. Lemma 4.1 shows that tolls satisfying the conditions of Lemma 3.2 induce an ordering on the marginal costs of the paths in the network.

**Lemma 4.1:** Any Nash flow \(f^{\text{nf}}\) induced by tolls according to Lemma 3.2 induces an ordering of the paths of \(G\) for which \(b_i < b_{i+1}\). Let \(n\) denote the number of paths with positive flow in \(f^{\text{nf}}\). Then \(f^{\text{nf}}\) satisfies the following set of equations for all \(i \in \{1, \ldots, n-1\}:

$$a_i f_i - a_{i+1} f_{i+1} = (b_{i+1} - b_i) z_i$$

where \(z_i \geq z_{i+1}, z_1 \leq 1, \) and \(z_{n-1} \geq 1/2\). Furthermore, \(2a_i f_i + b_i \geq 2a_{i+1} f_{i+1} + b_{i+1}\), or the first \(n\) elements of the vector \((2A f^{\text{nf}} + b)\) are ordered nonincreasing.

**Proof:** By the argument outlined in the proof of Lemma 3.2, if we define the user-specific coefficient \(\gamma_x = \frac{x_0(\kappa_1 - \kappa_2)}{1 + x_0(\kappa_1 - \kappa_2)}\), we can write the cost to user \(x\) of edge \(e\) as

$$J_e^x(f_e) = a_e f_e + b_e + \gamma_e a_e f_e,$$

where by hypothesis \(\gamma_x \in [0, 1]\) for all \(x\). Since \(\gamma : [0, r] \to [0, 1]\) is well-defined for any population \(s\) and tolling coefficients \(\kappa_1\) and \(\kappa_2\), we will henceforth simply speak of Nash flows induced by \(\gamma\).

The following arguments are an abridged version of those appearing in [4]; we include them here for completeness.

For any two paths \(p_i\) and \(p_{i+1}\), by hypothesis, \(b_i \leq b_{i+1}\). Consider a Nash flow \(f^{\text{nf}}\) induced by some \(\gamma\); suppose that in this flow, \(n\) paths have strictly positive flow. Take any user \(x \in [0, r]\) that is using path \(p_i+1\). Then since this is a Nash flow, user \(x\) must prefer path \(p_i+1\) to path \(p_i\), or

$$(1 + \gamma_x)(a_i f_i^{\text{nf}} - a_{i+1} f_{i+1}^{\text{nf}}) \geq b_{i+1} - b_i \geq 0.$$

Thus, \(a_i f_i^{\text{nf}} \geq a_{i+1} f_{i+1}^{\text{nf}} \geq 0\), for all \(i\); since \(\gamma_x \leq 1\), this further implies an ordering of marginal-costs; i.e., \(2a_i f_i^{\text{nf}} + b_i \geq 2a_{i+1} f_{i+1}^{\text{nf}} + b_{i+1}\).

Now, note that the left inequality in (31) is valid only for \(\gamma_x\) values greater than some threshold \(\gamma_i\); indicating that no user with \(\gamma_x > \gamma_i\) is using path \(p_i\). Similarly, no user with \(\gamma_x < \gamma_i\) is using path \(p_{i+1}\). That is, a higher sensitivity agent prefers higher-index paths. See [4] for an extended discussion of this toll-induced path-ordering.

Now, define \(z_i\) by the following:

$$z_i = \frac{a_i f_i^{\text{nf}} - a_{i+1} f_{i+1}^{\text{nf}}}{b_{i+1} - b_i}.$$ (32)

Note that for any \(f^{\text{nf}}\), this expression uniquely defines a set \(\{z_i\}_{i=1}^{n-1}\) of numbers. As a consequence of the ordering of \(\gamma\)-values on the paths of the network, these numbers satisfy \(z_i \geq z_{i+1}, z_1 \leq 1, \) and \(z_{n-1} \geq 1/2\).

Now, we introduce the following definition:

**Definition 1:** The derivative of \(f\) with respect to \(z_i\), denoted \(\frac{df}{dz_i} \in \mathbb{R}^n\), is the unique solution to the following:

3If any two paths \(p_i\) have \(b_i = b_j\), without loss of generality they may be combined into a single path \(p_k\) with \(a_k = (a_i + a_j)/(a_i + a_j)\) and \(b_k = b_i\).

4Uniqueness of \(\frac{df}{dz_i}\) is due to the fact that in a Nash flow, among paths with positive flow, at most one \(a_i\) coefficient can be zero.
Correspondingly, \( z_\delta < f \) allows us to complete the proof of the theorem. Index paths to low-index paths (that is, from paths with low yields a valid Nash flow in which traffic shifts from high- 
for some population. Let \( \delta \). Since all elements of \( \partial f / \partial z_i \) satisfy

\[
\sum_{j=1}^{n-1} \frac{\partial f_j}{\partial z_i} = 0.
\]

Now, Lemma 4.2 describes the relationship between \( \partial f \) and a Nash flow from which it was derived. Ultimately, Lemma 4.2 will allow us to show that under weakly-robust tolls, decreasing any agent’s sensitivity will increase the total latency of the resulting Nash flow.

**Lemma 4.2:** For any Nash flow \( f^n \) specify its corresponding \( \{z_i\}_{i=1}^{n-1} \) according to Lemma 4.1. If \( z_i < \min \{1, z_{i-1}\} \), then it is true that \( f^n + \delta \partial f^n / \partial z_i \) is also a Nash flow (for some different population) provided that \( \delta \) satisfies \( z \leq \min \{z_1, z_{i-1}\} - z_i, -f^n \cdot (\partial f^n / \partial z_i)^{-1} \). Furthermore, for all \( j \leq i \),

\[
\frac{\partial f^n}{\partial z_i} \geq 0,
\]

and for all \( k \geq i + 1 \),

\[
\frac{\partial f^n}{\partial z_i} \leq 0.
\]

Simply put, Lemma 4.2 states that a small increase in a single element of \( z \) (that is, a decrease in the associated \( \gamma \) value) yields a valid Nash flow in which traffic shifts from high-index paths to low-index paths (that is, from paths with low marginal-cost to paths with high marginal-cost). This crucial fact allows us to complete the proof of the theorem.

**Proof:** First we show the partitioning of \( \partial f^n / \partial z_i \). Note that Definition 1 implies that for all \( j \neq i \), \( \partial f^n / \partial z_i \) satisfies

\[
a_j \frac{\partial f_j}{\partial z_i} = a_{j+1} \frac{\partial f_{j+1}}{\partial z_i},
\]

and that for \( k \geq (i + 1) \),

\[
a_k \frac{\partial f_k}{\partial z_i} > a_k \frac{\partial f_k}{\partial z_i}.
\]

Since all elements of \( \partial f^n / \partial z_i \) sum to 0, the fact that all \( a_k > 0 \) except possibly \( a_k \) implies (33) and (34). A second consequence of the elements of \( \partial f^n / \partial z_i \) summing to 0 is that for small \( \delta < 0 \), \( f^n + \delta \partial f^n / \partial z_i \) represents a feasible flow on the network. Now we show that in fact, provided that we specify \( \delta \) carefully, \( f^n + \delta \partial f^n / \partial z_i \) represents a Nash flow for some population. Let \( \delta_z = \min \{\delta, z_{i-1}\} - z_i \), and \( \delta_f = -f^n \cdot (\partial f^n / \partial z_i)^{-1} \). Note that \( \delta_z \) is simply “how much room” \( z_i \) has to increase until it equals the lesser of 1 and \( z_{i-1} \). Correspondingly, \( \delta_f \) is the threshold at which \( f^n + \delta \partial f^n / \partial z_i \) has exactly zero flow on path \( p_i \). Thus, for \( \delta = \min \{\delta_z, \delta_f\} \), it is always true that \( f^n + \delta \partial f^n / \partial z_i \) is a feasible flow which specifies a collection of \( z \)-values according to the definition in Lemma 4.1.

Now, note that any feasible flow which satisfies (32) for descending-ordered \( z_i \) can be described as a Nash flow induced by some \( \gamma \); specific values of \( \gamma \) can be computed by letting \( \gamma_i = \frac{1}{z_i} - 1 \). That is, \( f^n + \delta \partial f^n / \partial z_i \) represents a Nash flow induced by some \( \gamma \).

**Proof of Theorem 3.3:** Let \( f^n \) be a Nash flow for some heterogeneous population. We shall construct a sequence of Nash flows \( f(k) \) satisfying \( L(f(k+1)) \geq L(f(k)) \) for which \( f(0) = f^n \) and \( f(T) \) is an un-tolled Nash flow, for some finite \( T \).

We construct the sequence iteratively as follows: For any \( f(k) \), compute the \( z \)-vector associated with \( f^n \) according to Lemma 4.1. Select \( i \) as the lowest index satisfying \( z_i < \min \{1, z_{i-1}\} \). If no element of \( z \) satisfies these conditions, it must be true that \( z = 1_{k-1} \) and \( f(k) \) describes an un-tolled Nash flow, so the iteration terminates. Let \( \delta_z = \min \{1, z_{i-1}\} - z_i \) and \( \delta_f = -f^n(k) \cdot (\partial f^n(k) / \partial z_i)^{-1} \). Define the successor flow \( f(k+1) \) by

\[
f(k+1) = f(k) + \min \{\delta_z, \delta_f\} \cdot \frac{\delta f(k)}{\partial z_i}.
\]

If \( f(k) \) is a Nash flow, then Lemma 4.2 guarantees that \( f(k+1) \) is as well.

To see that \( L(f(k+1)) \geq L(f(k)) \), note that by Lemma 4.1, \( 2Af^n + b \) is ordered nonincreasing. By Lemma 4.2, \( \partial f^n / \partial z_i \) is partitioned such that elements \( 1 \) through \( i \) are nonnegative, and elements \( i+1 \) through \( n \) are nonpositive, and \( 1^T (\partial f^n / \partial z_i) = 0 \). Thus, the derivative of \( L \) along the line segment \( f(k+1) - f(k) \) is always nonnegative:

\[
\frac{\partial L(f^n)}{\partial z_i} = \frac{\partial f^n}{\partial z_i} \cdot (2Af^n + b) \geq 0.
\]

The iteration must always terminate at an un-tolled Nash flow in at most \( 2n - 2 \) steps; \( n - 1 \) steps for equalizing the elements of \( z \), and \( n - 1 \) steps for “zeroing” the flow on a network path.

Thus, since \( L(f^n) \leq L(f^{untolled}) \) for an untolled Nash flow \( f^{untolled} \), we have shown that the necessary conditions of Lemma 3.2 are also sufficient for a taxation mechanism’s weak robustness on parallel networks.

Finally, note that if the iteration defined above is modified slightly so that \( z_1 \) is held constant and all other \( z_i \) values are increased to \( z_1 \) (rather than to 1), the resulting sequence \( L(f(k)) \) is still monotone-increasing, but now terminates at a Nash flow resulting associated with a homogeneous population with sensitivity \( S_L \). That is, the worst-case Nash flows for any weakly-robust tolls are caused by extreme low-sensitivity homogeneous populations; thus, congestion-minimizing tolls should be designed to maximize \( S_L(S_1 S_2 \cdots S_n) \) subject to \( S_L \geq 1 \). Given this fact, it is easy to compute (18), and (19) follows by taking the limit as \( S_U \to +\infty \).