The Benefit of *Perversity* in Taxation Mechanisms for Distributed Routing

Philip N. Brown and Jason R. Marden

Abstract— **We study pricing as a means to improve the congestion experienced by self-interested traffic. When user price-sensitivities are unknown, it is not generally possible to incentivize optimal flows with static pricing. Nonetheless, recent work has derived non-trivial pricing that can be guaranteed never to be *perverse*; that is, for every network and every user population, the routing incentivized by these prices is never worse than un-influenced routing. In this paper, we ask two sides of a dual question: first, if tolls *were* allowed to create perverse incentives on some networks, could this improve worst-case outcomes overall? Second, how perverse are the tolls that minimize congestion in worst-case over all networks? Perhaps counterintuitively, we show that if a system operator is willing to risk creating small perversities on some networks, this can reduce worst-case congestion. Subsequently, we derive the tolls that minimize worst-case congestion over all parallel networks, and show that they always outperform the best non-perverse tolls. In other words, there is a tradeoff between perversity and efficiency: even on parallel networks, a taxation mechanism that optimizes worst-case performance does so by creating perverse incentives on some networks.**

I. INTRODUCTION

For at least a century, economists have suggested pricing public resources as a means to encourage their efficient utilization [1]. As engineered systems become ever-more interconnected with their users, the opportunities for and potential benefits of smart pricing schemes are becoming increasingly apparent. Accordingly, research is being conducted on pricing as a means to influence human behavior in a wide range of applications, including urban traffic congestion, ridesharing platforms, wireless networks, auctions, and others [2]–[4].

One such socio-technical system is a transportation network, in which commuters make local routing decisions in response to personal objectives – objectives which may be misaligned with those of the system operator. A common model for the transportation network setting is the *non-atomic congestion game*, in which traffic is modeled as a continuum of agents, each infinitesimally small and individually having no impact on congestion. The un-influenced social state in the game is known as a *Nash flow* (or *Wardrop equilibrium*), defined as a routing profile in which no driver can switch paths and obtain a lower cost [5].

In settings like these, it is known that social systems can perform "poorly," from the standpoint that individual decision-making may lead the aggregate social behavior to a suboptimal state. For the transportation network, this means that if a system planner could explicitly prescribe a route for each commuter, it is possible to decrease congestion costs significantly when compared to the un-influenced case. This inefficiency due to self-interested decisions is termed the *price of anarchy*, formally defined as the ratio between the total congestion of a Nash flow and that of an optimal flow, taken in worst case over a class of games [6]. It is known that if all delay functions are linear, the price of anarchy in network routing can be as high as $4/3$ [7].

Much work has focused on tolls as a means to reduce the price of anarchy; seminal work indicated that given a perfect characterization of a routing problem (e.g., network topology, user tax-sensitivities, demand), a system operator can levy tolls which induce perfectly-optimal flows [8]–[10]. Recently, interest has been growing in questions of robustness: how should a designer choose prices when the underlying system is uncertain or time-varying? Work along these lines has prescribed solutions for unknown or time-varying demand [11]–[13], unknown latency functions [14], and unknown networks and tax-sensitivities [2].

As the price of anarchy is a worst-case metric, it may happen that pricing that minimizes price of anarchy for a class of problems need not perform well on *particular* problems in that class. That is, optimizing worst-case performance could very well come at the expense of typical-case performance. This is the purview of this paper: we seek to study specifically how much is lost on individual problem instances when minimizing the price of anarchy.

We study the relationship between the price of anarchy and a new robustness metric that we formally introduce in this paper termed *perversity*. The perversity of a taxation mechanism is defined as the ratio between the performance it incentivizes and the un-influenced performance, taken in worst case over user populations and networks. Thus, if a taxation mechanism has a perversity of $\alpha \geq 1$, this indicates that for *every* network and user population, it incentivizes Nash flows that are never more than an α factor more costly than the un-tolled Nash flows. Preliminary work has shown that when user populations are diverse, a non-trivial taxation mechanism must rely in some way on network topology if it is to have a perversity of 1 [15]. Nevertheless, it is known that the class of parallel networks *does* admit taxation mechanisms with a perversity of 1. This concept is related to recent observations that a utility design which minimizes worst-case performance does so at the expense of best-case performance [16].

Accordingly, following recent work in [15], [17] that stud-

ies locally-computed incentives, we investigate the tradeoff between the price of anarchy and perversity. Specifically, we ask the following questions:

1) If a system operator is willing to risk creating small perversities, can this improve the price of anarchy?
2) How perverse are the tolls that optimize the price of anarchy?
3) How can a designer explicitly enforce a constraint on the degree of perversity?

We ask these questions for parallel-network, affine-cost routing games; we choose this simple class of games for two reasons: first, the class of parallel networks is the largest class known that admits a non-trivial locally-computed taxation mechanism that is guaranteed to create no perverse incentives [15]. Second, the restriction to affine costs allows us to apply tools from [2] to explicitly characterize the Nash flows resulting from tolls. Furthermore, we investigate a simple class of tolls: the scaled marginal-cost taxation mechanism. It is known that all locally-computed non-perverse taxation mechanisms,[1] are a generalization of the scaled marginal-cost taxation mechanism, so many of the salient features of these mechanisms can be studied by our analysis here [15].

Our results are summarized as follows: We show in Theorem 4.1 that the optimal non-perverse mechanism from [15] *does not* minimize the price of anarchy over all parallel networks. That is, in some settings, a particular aggressive mechanism reported in [2] provides a significantly better price of anarchy than any non-perverse mechanism. Figure 2 shows this clearly: the dotted blue line indicates the price of anarchy of the non-perverse mechanism from [15], whereas the aggressive mechanism from [2] is shown to have a lower price of anarchy for large values of S_L/S_U.

Next, we show that to minimize the price of anarchy, a perversity strictly greater than 1 is *required*. We provide plots of the price of anarchy for various parameters, and prove that the price of anarchy of the best non-perverse mechanism is strictly worse than the best-achievable price of anarchy over all mechanisms, showing that gains in the price of anarchy can only come at the risk of creating perverse incentives.

Finally, we explicitly characterize the marginal-cost tolls that cause no more than an $\alpha \geq 1$ degradation in the performance of any Nash flow. This can be viewed as a safety constraint: perhaps a designer seeks tolls that minimize the price of anarchy for a relatively well-known setting, but wants to hedge against major inefficiencies caused by some other unmodeled adverse conditions.

Figure 1 illustrates this last concept graphically. Here, the three lines correspond to three different uncertainty (or variance) environments; higher price of anarchy corresponds to higher uncertainty. The horizontal axis represents the system operator's tolerance for perversity; note that if the operator does not allow perversity (the left side of the plots), the price of anarchy is somewhat higher than if the operator can permit some degree of perversity (the right side of the plots). We point out that when perversity is 1, this
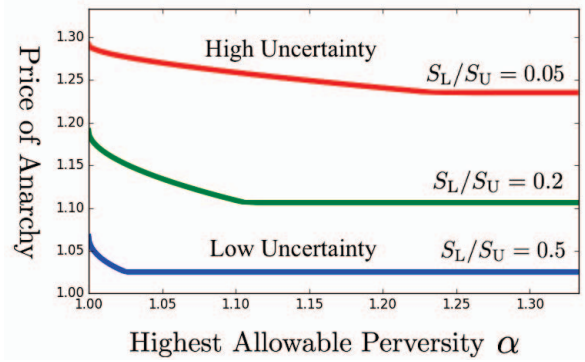
Fig. 1. Price of anarchy with respect to limits on perversity. Each plot represents the best-achievable price of anarchy for homogeneous populations on parallel networks as a function of the highest-allowable perversity of scaled marginal-cost tolls (α in Theorem 4.3). The three lines are for three different values of S_L/S_U, which are a crude meausre of a designer's uncertainty about the population: low S_L/S_U means high uncertainty (or high variance). For example, the green dashed line (for $S_L/S_U = 0.2$) shows that if the $\alpha = 1$ (i.e., perversity is not allowed), the price of anarchy is just under 1.2; if α is allowed to be greater than about 1.1, then the price of anarchy can be nearly as low as 1.1.

corresponds to the non-perverse setting studied in [15] and the price of anarchy we report in Theorem 4.1. On the other hand, the right side of the plots (essentially corresponding to unconstrained perversity) corresponds to the best-achievable price of anarchy, which is the subject of Theorem 4.2.

II. MODEL AND PERFORMANCE METRICS

A. Routing Game

Consider a network routing problem for a network (V, E) comprised of vertex set V and edge set E. A mass of r units of traffic needs to be routed from a common source $\sigma \in V$ to a common destination $t \in V$. We write $\mathcal{P}$ to denote the set of *paths* available to the traffic, where each path $p \in \mathcal{P}$ consists of a set of edges connecting s to t. Note that this paper considers only the case of *symmetric* (or single-commodity) routing problems, in which all traffic can access the same set of paths. 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 $\sum_{p \in \mathcal{P}} f_p = r$, where $f_p \geq 0$ denotes the mass of traffic on path p. 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 affine latency function of the form $\ell_e(f_e) = a_e f_e + b_e$, where $\ell_e(f_e)$ denotes the delay experienced by users of edge e when the edge flow is f_e, and $a_e \geq 0$ and $b_e \geq 0$ are edge-specific constants. We measure the cost of a flow f by its *total latency*, given by

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

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

$$f^* \in \underset{f \text{ is feasible}}{\operatorname{argmin}} \ \mathcal{L}(f). \quad (2)$$

A *routing problem* is given by $G = (V, E, r, \{\ell_e\})$. We denote by $\mathcal{G}^p$ the set of all parallel-path routing problems with affine latency functions.

To study the effect of taxes on self-interested behavior, we model the above routing problem as a heterogeneous non-atomic congestion game. We assign each edge $e \in E$ a flow-dependent taxation function $\tau_e : \mathbb{R}^+ \to \mathbb{R}^+$. To characterize users' taxation sensitivities, let each user $x \in [0, r]$ have a taxation sensitivity $s_x \in [S_\mathrm{L}, S_\mathrm{U}] \subseteq \mathbb{R}^+$, where $S_\mathrm{L} \geq 0$ and $S_\mathrm{U} < \infty$ are lower and upper sensitivity bounds, respectively. We capture a specification of sensitivities with a function $s : [0, r] \to [S_\mathrm{L}, S_\mathrm{U}]$ (where, for notational compactness, s evaluated at x is written s_x). If all users in s have the same sensitivity (i.e., $s_x = s_y$ for all $x, y \in [0, r]$), the population is said to be *homogeneous*; otherwise it is *heterogeneous*. We write the extreme-sensitivity homogeneous populations with sensitivities S_L and S_U as s^L and s^U, respectively. The set of possible sensitivity distributions is the set of Lebesgue-measurable functions $\mathcal{S} = \{s : [0, r] \to [S_\mathrm{L}, S_\mathrm{U}]\}$.

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

$$J_p^x(f) = \sum_{e \in p} [\ell_e(f_e) + s_x \tau_e(f_e)], \qquad (3)$$

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 all users are individually using minimum-cost paths given the choices of other users, or if for all users $x \in [0, r]$ we have

$$J^x(f) = \min_{p' \in \mathcal{P}} \sum_{e \in p'} [\ell_e(f_e) + s_x \tau_e(f_e)]. \qquad (4)$$

It is well-known that a Nash flow exists for any non-atomic congestion game of the above form [5]. On parallel-path networks, for any (even heterogeneous) population s, these Nash flows are unique in cost (that is, all Nash flows have the same total latency) [18].

B. Taxation Mechanisms and Performance Metrics

To model locally-computed tolls, we consider so-called *network-agnostic* taxation mechanisms as in [17]. Here, each edge's taxation function is computed using only locally-available information. That is, $\tau_e(f_e)$ depends only on ℓ_e, not on edge e's location in the network, the network topology, the overall traffic rate, or the properties of any other edge. A network-agnostic taxation mechanism T is thus a mapping from latency functions to taxation functions, and the taxation function associated with latency function ℓ_e is given by

$$\tau_e(\cdot) = T(\ell_e). \qquad (5)$$

To evaluate the performance of taxation mechanisms, we write $\mathcal{L}^\mathrm{nf}(G, s, T)$ to denote the total latency of a Nash flow for routing problem G and population s induced by taxation mechanism T. Let $\mathcal{L}^*(G)$ denote the total latency of the optimal flow on G. The *price of anarchy* compares the Nash flows induced by taxes with the optimal flows; here, we define the price of anarchy of a class of games $\mathcal{G}$ under the

influence of taxes T as the following measure, in worst-case over networks and user populations:

$$\mathrm{PoA}(\mathcal{G}, T) \triangleq \sup_{G \in \mathcal{G}} \sup_{s \in \mathcal{S}} \frac{\mathcal{L}^\mathrm{nf}(G, s, T)}{\mathcal{L}^*(G)}. \qquad (6)$$

In this paper, we pose a somewhat different question: rather than measuring how far the influenced flows are from *optimal*, it may be desirable to measure how much the taxes help *with respect to the un-influenced flows*. Here, we may define a similar metric to that in (6), but with the un-influenced total latency $\mathcal{L}^\mathrm{nf}(G, \emptyset)$ in the denominator. We call such a metric the *index of perversity* of taxation mechanism T, defined as

$$\mathrm{PI}(\mathcal{G}, T) \triangleq \sup_{G \in \mathcal{G}} \sup_{s \in \mathcal{S}} \frac{\mathcal{L}^\mathrm{nf}(G, s, T)}{\mathcal{L}^\mathrm{nf}(G, \emptyset)}. \qquad (7)$$

Here, if T has a large index of perversity, this means that on some networks, it incentivizes flows that are considerably worse than the un-influenced flows; in other words, it can create perverse incentives. It has recently been shown that the index of perversity of conservatively-scaled marginal-cost tolls on parallel networks is 1; that is, these tolls never create perverse incentives [15].

Note that it is always true that $\mathrm{PI}(\mathcal{G}, T) \leq \mathrm{PoA}(\mathcal{G}, T)$; this is because on any G, $\mathcal{L}^\mathrm{nf}(G, \emptyset) \geq \mathcal{L}^*(G)$. Finally, when these metrics are evaluated only over homogeneous populations (as opposed to heterogeneous), we write them as $\mathrm{PoA}^\mathrm{hm}(\mathcal{G}, T)$ and $\mathrm{PI}^\mathrm{hm}(\mathcal{G}, T)$, respectively.

III. RELATED WORK

The classical example of a network-agnostic taxation mechanism is that of the *marginal-cost* or *Pigovian* taxation mechanism T^mc. For any edge e with latency function $\ell_e(f_e) = a_e f_e + b_e$, the accompanying marginal-cost toll is

$$\tau_e^\mathrm{mc}(f_e) = a_e f_e, \quad \forall f_e \geq 0, \qquad (8)$$

In [19] the authors show that for any G, it is true that $\mathcal{L}^*(G) = \mathcal{L}^\mathrm{nf}(G, s, T^\mathrm{mc})$, provided that s is a homogeneous population with sensitivity equal to 1.

Recent research has shown that many salient features of network-agnostic taxation mechanisms can be analyzed by studying scaled versions of marginal-cost tolls [15]. Accordingly, in this paper, we investigate various scaled marginal cost tolls. The first of these is the subject of [2]. There, it is shown that for a subset of parallel networks,[2] the scaled marginal-cost tolls which minimize the price of anarchy are the "aggressive" taxation mechanism T^AGG assigning tolls of

$$\tau_e^\mathrm{AGG}(f_e) = \frac{a_e f_e}{\sqrt{S_\mathrm{L} S_\mathrm{U}}}. \qquad (9)$$

The tolls in (9) are aggressive in this sense: any user with sensitivity S_U experiences T^AGG a factor of $\sqrt{S_\mathrm{U}/S_\mathrm{L}}$ more than the nominal marginal-cost toll (8); these users are essentially being over-charged. It is shown in [15] that

---

[2]In [2], the considered networks are those with flow on all edges in an un-tolled Nash flow. Under this assumption, *all* scaled marginal-cost tolls have an index of perversity of 1. In this paper, we relax this assumption.

this over-charging can create perverse incentives, and that the optimal scaled marginal-cost tolls that *do not* create perverse incentives are the "conservative" tolls T^{CON}, which for any edge e, assigns a taxation function of

$$\tau_e^{\mathrm{CON}}(f_e) = \frac{a_e f_e}{S_{\mathrm{U}}}. \tag{10}$$

By charging every agent as though they had sensitivity S_{U}, these tolls ensure that no agent is over-charged, which ensures that $\mathrm{PI}(\mathcal{G}^p, T^{\mathrm{CON}}) = 1$ Furthermore, of all scaled marginal-cost tolls with index of perversity 1, T^{CON} minimizes the price of anarchy.

IV. OUR CONTRIBUTIONS

In this paper we investigate some fundamental tradeoffs between the price of anarchy and perversity, and show that in some cases, allowing perverse incentives can modestly *decrease* the price of anarchy.

A. The endpoints: aggressive versus conservative tolls

We begin with the following theorem, which reports the price of anarchy of T^{AGG} and T^{CON}, and shows that T^{AGG} creates perverse incentives on some networks.

Theorem 4.1: For heterogeneous populations, the following are true:

$$\mathrm{PoA}\left(\mathcal{G}^p, T^{\mathrm{CON}}\right) = \frac{4}{4\left(1 + \frac{S_{\mathrm{L}}}{S_{\mathrm{U}}}\right) - \left(1 + \frac{S_{\mathrm{L}}}{S_{\mathrm{U}}}\right)^2}. \tag{11}$$

$$\mathrm{PoA}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right) = \frac{1}{2} + \frac{1}{4}\left(\sqrt{\frac{S_{\mathrm{U}}}{S_{\mathrm{L}}}} + \sqrt{\frac{S_{\mathrm{L}}}{S_{\mathrm{U}}}}\right). \tag{12}$$

Furthermore, under T^{AGG}, the index of perversity is equal to the price of anarchy:

$$\mathrm{PI}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right) = \mathrm{PoA}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right). \tag{13}$$

These are depicted graphically in Figure 2. The horizontal axis in Figure 2 is the sensitivity ratio $S_{\mathrm{L}}/S_{\mathrm{U}}$, which can be viewed as a proxy for the precision with which the user population is characterized. When $S_{\mathrm{L}}/S_{\mathrm{U}} \to 0$, populations are essentially totally unknown, so the price of anarchy tends to be large. When $S_{\mathrm{L}}/S_{\mathrm{U}} \to 1$, populations are both well-known and homogeneous, so the price of anarchy tends to 1. Note that for poorly-characterized populations, $\mathrm{PoA}(\mathcal{G}^p, T^{\mathrm{AGG}})$ is quite large (tending to infinity as $S_{\mathrm{L}}/S_{\mathrm{U}} \to 0$), indicating that aggressively-scaled tolls carry a risk of incentivizing highly pathological Nash flows.

On the other hand, for relatively well-characterized populations (approximately for $S_{\mathrm{L}}/S_{\mathrm{U}} > 0.17$), the aggressive tolls yield a lower price of anarchy than the conservative tolls. This illustrates one of the central points of this paper: tolls that systematically avoid perverse incentives cannot, in general, be optimal with respect to the price of anarchy. Equation (13) indicates that even when aggressive tolls have a low price of anarchy, they can only achieve this at the expense of causing perverse incentives on some networks.
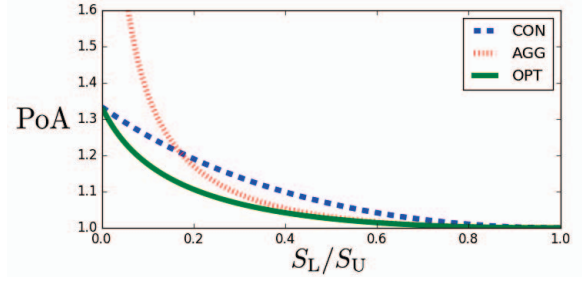

Fig. 2. Price of anarchy plot for Theorem 4.1. The dotted red line depicts the price of anarchy for the aggressive T^{AGG}, which charges marginal-cost tolls scaled by $(S_{\mathrm{L}}S_{\mathrm{U}})^{-1/2}$. The dashed blue line depicts the price of anarchy for the conservative T^{CON}, which charges marginal-cost tolls scaled by $1/S_{\mathrm{U}}$. The solid green line depicts the (homogeneous) price of anarchy for the optimal $T(\kappa^*)$ defined in Theorem 4.2. This illustrates the principle that aggressive behavior-influencing mechanisms may be warranted when the uncertainty is low, but when uncertainty is high, these carry the risk of incentivizing severe inefficiencies.

Furthermore, this illustrates the intuitive idea that high uncertainty necessitates conservatism, but that as uncertainty decreases, more-aggressive approaches may be warranted.

Proof: First consider T^{CON}. This taxation mechanism is studied in [15, Lemma 5.2], where it is shown for any parallel-network routing problem $G \in \mathcal{G}^p$ and any heterogeneous $s \in \mathcal{S}$ that

$$\mathcal{L}^{\mathrm{nf}}\left(G, s, T^{\mathrm{CON}}\right) \leq \mathcal{L}^{\mathrm{nf}}\left(G, s^{\mathrm{L}}, T^{\mathrm{CON}}\right), \tag{14}$$

where s^{L} denotes a homogeneous population with sensitivity S_{L}. Thus, the price of anarchy here reduces to studying the worst-case performance of s^{L}:

$$\mathrm{PoA}\left(\mathcal{G}^p, T^{\mathrm{CON}}\right) = \sup_{G \in \mathcal{G}} \frac{\mathcal{L}^{\mathrm{nf}}\left(G, s^{\mathrm{L}}, T^{\mathrm{CON}}\right)}{\mathcal{L}^*(G)}. \tag{15}$$

We can then directly apply Lemma 5.1 (see Appendix), which relates the price of anarchy to a stylized homogeneous price-sensitivity β ($\beta > 1$ means users are overcharged, while $\beta < 1$ means they are under-charged). The conservative taxation mechanism T^{CON} assigns taxes of $a_e f_e / S_{\mathrm{U}}$, so a population with sensitivity S_{L} has a value of $\beta = S_{\mathrm{L}}/S_{\mathrm{U}} \leq 1$, which implies (11).

To show (12), we appeal to Lemma 5.2 (see Appendix), which states that

$$\mathcal{L}^{\mathrm{nf}}\left(G, r, s, T^{\mathrm{AGG}}\right) \leq \mathcal{L}^{\mathrm{nf}}\left(G, r, s^{\mathrm{U}}, T^{\mathrm{AGG}}\right). \tag{16}$$

That is, in the case of T^{AGG}, the price of anarchy reduces to studying worst-case performance of homogeneous populations with sensitivity S_{U}:

$$\mathrm{PoA}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right) = \sup_{G \in \mathcal{G}} \frac{\mathcal{L}^{\mathrm{nf}}\left(G, s^{\mathrm{U}}, T^{\mathrm{AGG}}\right)}{\mathcal{L}^*(G)}. \tag{17}$$

As before, we can then directly apply (29) from Lemma 5.1. In our case, T^{AGG} assigns taxes of $a_e f_e / \sqrt{S_{\mathrm{L}} S_{\mathrm{U}}}$, so a population with sensitivity S_{U} has a value of $\beta = \sqrt{S_{\mathrm{U}}/S_{\mathrm{L}}} \geq 1$, which implies (12).

To show the final point, recall that it is always true that $\mathrm{PI}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right) \leq \mathrm{PoA}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right)$. Thus, to show (13), we must construct a matching lower bound on $\mathrm{PI}\left(\mathcal{G}^p, T^{\mathrm{AGG}}\right)$

for any S_U. To this end, consider a two-link network with latency functions $\ell_1(f_1) = a_1 f_1$, where $a_1 = 2\sqrt{S_U/S_L}/(1 + \sqrt{S_U/S_L})^2$, and $\ell_2(f_2) = 1$. Let $r = 1$, and note that the un-tolled flow $f^\emptyset$ has $f_1^\emptyset = 1$, so that

$$\mathcal{L}^{\mathrm{nf}}(G, \emptyset) = \frac{2\sqrt{S_U/S_L}}{\left(1 + \sqrt{S_U/S_L}\right)^2}.$$

On the other hand, for agents with sensitivity $s = S_U$, T^{AGG} induces a flow f^τ with

$$f_1^\tau = \frac{\sqrt{S_U/S_L} + 1}{2\sqrt{S_U/S_L}} \quad \text{and} \quad f_2^\tau = \frac{\sqrt{S_U/S_L} - 1}{2\sqrt{S_U/S_L}}. \quad (18)$$

This is the case because $1 + \kappa S_U = 1 + \sqrt{S_U/S_L}$, so

$$(1 + \kappa S_U) a_1 f_1^\tau = \frac{2\sqrt{S_U/S_L}}{\left(1 + \sqrt{S_U/S_L}\right)^2} \cdot \frac{\left(\sqrt{S_U/S_L} + 1\right)^2}{2\sqrt{S_U/S_L}} \quad (19)$$

$$= 1 = \ell_2(f_2^\tau). \quad (20)$$

At this flow the total latency can be verified to be

$$\mathcal{L}^{\mathrm{nf}}\left(G, s^U, T^{\mathrm{AGG}}\right) = \frac{1}{2}. \quad (21)$$

Thus, $\mathcal{L}^{\mathrm{nf}}\left(G, s^U, T^{\mathrm{AGG}}\right)/\mathcal{L}^{\mathrm{nf}}(G, \emptyset)$ is equal to (12). $\blacksquare$

B. Optimal tolls must create perverse incentives

Theorem 4.1 and its associated plots in Figure 2 demonstrate that neither the aggressive T^{AGG} nor the conservative T^{CON} minimize the price of anarchy for all situations. In this section we ask what taxation mechanism *does* minimize the price of anarchy for a given S_L and S_U. Here, we study only the price of anarchy for homogeneous populations. Theorem 4.2 shows that this optimal taxation mechanism is strictly "more aggressive" than T^{CON}, in the sense that it charges more than T^{CON} on every link. Intuitively, this is the taxation mechanism that perfectly balances the harm that can be caused by a homogeneous S_L population with the harm that can be caused by a homogeneous S_U population.

In the following, let $T(\kappa)$ denote the scaled marginal-cost taxation mechanism assigning taxation functions of $\tau_e(f_e) = \kappa a_e f_e$ to each network edge.

Theorem 4.2: For homogeneous populations, for any $0 \le S_L < S_U$, the PoA-minimizing toll scalar

$$\kappa^* \triangleq \inf_{\kappa \ge 0} \mathrm{PoA}^{\mathrm{hm}}\left(\mathcal{G}^p, T(\kappa)\right) \quad (22)$$

is the unique solution κ^* on the interval $(1/S_U, 1/S_L)$ to

$$\frac{4}{4\left(1 + \kappa^* S_L\right) - \left(1 + \kappa^* S_L\right)^2} = \frac{\left(1 + \kappa^* S_U\right)^2}{4\kappa^* S_U}. \quad (23)$$

Its resulting price of anarchy and perversity index for homogeneous populations are equal and greater than one:

$$\mathrm{PoA}^{\mathrm{hm}}(\mathcal{G}^p, T(\kappa^*)) = \mathrm{PI}^{\mathrm{hm}}(\mathcal{G}^p, T(\kappa^*)) = \frac{\left(1 + \kappa^* S_U\right)^2}{4\kappa^* S_U} > 1. \quad (24)$$

The resulting price of anarchy and perversity index is plotted in Figure 2 along with the results from Theorem 4.1.

Proof: The restriction to homogeneous populations here allows us to apply Lemma 5.1 directly, and leverage its monotonicity properties to obtain the result. Consider the expressions given by (28) and (29). Here, for a given $\kappa \ge 0$, β is simply given by κS_L or κS_U. Note that (28) is monotone decreasing on its domain and (29) is monotone increasing on its domain, and both are continuous for all $\beta \ge 0$. This monotonicity guarantees that for a fixed κ, the price of anarchy is realized by an extreme-sensitivity population:

$$\mathrm{PoA}^{\mathrm{hm}}(\mathcal{G}^p, T(\kappa^*)) = \max_{s \in \{s^L, s^U\}} \mathrm{PoA}^{\mathrm{hm}}(\mathcal{G}^p, s, T(\kappa^*)). \quad (25)$$

This also implies that the optimal κ^* must be on the interval $[1/S_U, 1/S_L]$. If, for example, κ is less than $1/S_U$, then the price of anarchy for *both* S_L and S_U populations can be decreased by raising κ to $1/S_U$.

Thus, for $\kappa \in [1/S_U, 1/S_L]$, the price of anarchy is either realized by S_L (generating a value of $\beta = \kappa S_L \le 1$) or by S_U (generating a value of $\beta = \kappa S_U \ge 1$). The S_L price of anarchy is 1 when $\kappa = 1/S_L$ and the S_U price of anarchy is 1 when $\kappa = 1/S_U$. Thus, since the functions are strictly increasing (or decreasing), the graphs of their induced price of anarchy as a function of κ must intersect at some point on the interval $(1/S_U, 1/S_L)$. This intersection is the solution to the equation described in (23), which specifies the price-of-anarchy-minimizing κ^*.

The fact that the perversity index is equal to the price of anarchy is due to the fact that the price of anarchy is realized by a population with sensitivity S_U when $\kappa > 1/S_U$ and can be shown with a simple two-link network. It is strictly greater than 1 because $\kappa^* > 1/S_U$. $\blacksquare$

C. Limiting the perversity of tolls

We have so far assumed that minimizing the price of anarchy is the system planner's sole design goal, and Theorem 4.2 has shown that this cannot be done without creating perverse incentives on some networks. Nonetheless, there may be situations in which the tolls must satisfy an additional upper bound on perversity. The present following theorem gives a simple constraint on κ in terms of α and S_U that guarantees an index of perversity no worse than α.

Theorem 4.3: For homogeneous populations, for any $\alpha \ge 1$, if

$$\kappa \le \frac{2\alpha - 1 + 2\sqrt{\alpha(\alpha - 1)}}{S_U}, \quad (26)$$

then

$$\mathrm{PI}^{\mathrm{hm}}\left(\mathcal{G}^p, \kappa\right) = \frac{\left(1 + \kappa S_U\right)^2}{4\kappa S_U} \le \alpha. \quad (27)$$

Proof: For homogeneous populations, whenever $\kappa \le S_L$, the index of perversity is realized by a S_U-sensitivity population. First, the index of perversity cannot be realized by a population with sensitivity $s \le 1/\kappa$; this is an implication of Theorem 4.1 of [15]. Second, it is simple to show on a two-link network that the worst-case perversity for populations with sensitivity $s \ge 1/\kappa$ is in fact realized by

populations with sensitivity S_{U}. Since the index of perversity is never more than the price of anarchy, we can simply use the price of anarchy expression from (29) in Lemma 5.1 to calculate the bound in (26) (namely, by making the inequality in (27) into an equality and solving for κ). ∎

V. Conclusion

This paper is part of an ongoing endeavor to understand the tradeoffs and consequences associated with influencing user behavior in transportation networks. Here, we have shown for a small class of games and a particular form of taxation mechanism that any taxes which minimize the price of anarchy do so at the expense of the perversity index. Naturally, it would be desirable to prove the bounds in Theorems 4.2 and 4.3 for the case of heterogeneous populations as well as homogeneous. Looking further ahead, it will be interesting to see how the answers to these questions change when taxation mechanisms are allowed to be network-dependent in some way.

References

[1] A. C. Pigou, *The Economics of Welfare*. Macmillan, New York, 1920.
[2] P. N. Brown and J. R. Marden, "The Robustness of Marginal-Cost Taxes in Affine Congestion Games," *Transactions on Automatic Control*, vol. 62, no. 8, pp. 3999–4004, 2017.
[3] S. Banerjee, R. Johari, and C. Riquelme, "Pricing in Ride-Sharing Platforms : A Queueing-Theoretic Approach," in *16th ACM Conference on Economics and Computation*, pp. 639–639, 2015.
[4] T. H. Noe, M. Rebello, and J. Wang, "Learning to bid: The design of auctions under uncertainty and adaptation," *Games and Economic Behavior*, vol. 74, pp. 620–636, mar 2012.
[5] J. Wardrop, "Some Theoretical Aspects of Road Traffic Research," in *Proc. of the Institution of Civil Engineers, Part II, Vol. 1, No. 36*, pp. 352–362, 1952.
[6] C. Papadimitriou, "Algorithms, games, and the internet," in *STOC '01: Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pp. 749–753, 2001.
[7] T. Roughgarden, *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.
[8] L. Fleischer, K. Jain, and M. Mahdian, "Tolls for Heterogeneous Selfish Users in Multicommodity Networks and Generalized Congestion Games," in *Proc. 45th IEEE Symp. on Foundations of Computer Science*, (Rome, Italy), pp. 277–285, 2004.
[9] G. Karakostas and S. Kolliopoulos, "Edge pricing of multicommodity networks for heterogeneous selfish users," in *Proc. 45th IEEE Symp. on Foundations of Computer Science*, (Rome, Italy), pp. 268–276, 2004.
[10] W. Sandholm, "Negative Externalities and Evolutionary Implementation," *The Review of Economic Studies*, vol. 72, pp. 885–915, jul 2005.
[11] H. Yang, Q. Meng, and D. H. Lee, "Trial-and-error implementation of marginal-cost pricing on networks in the absence of demand functions," *Transportation Research Part B: Methodological*, vol. 38, no. 6, pp. 477–493, 2004.
[12] F. Farokhi and K. H. Johansson, "A piecewise-constant congestion taxing policy for repeated routing games," *Transportation Research Part B: Methodological*, vol. 78, pp. 123–143, 2015.
[13] G. Christodoulou, K. Mehlhorn, and E. Pyrga, "Improving the price of Anarchy for Selfish Routing via coordination mechanisms," *Algorithmica*, vol. 69, no. 3, pp. 619–640, 2014.
[14] U. Bhaskar, K. Ligett, L. Schulman, and C. Swamy, "Achieving Target Equilibria in Network Routing Games without Knowing the Latency Functions," in *Proc. IEEE Symp. Foundations of Computer Science*, (Philadelphia, PA, USA), pp. 31–40, 2014.
[15] P. N. Brown and J. R. Marden, "Fundamental Limits of Locally-Computed Incentives in Network Routing," in *American Control Conf.*, (Seattle, WA, USA), pp. 5263–5268, 2017.
[16] V. Ramaswamy and J. R. Marden, "The Inherent Tension Between Anarchy and Stability in Noncooperative Designs," 2017.
[17] P. N. Brown and J. R. Marden, "Studies on Robust Social Influence Mechanisms: Incentives for Efficient Network Routing in Uncertain Settings," *IEEE Control Systems Magazine*, vol. 37, no. 1, pp. 98–115, 2017.
[18] I. Milchtaich, "Topological Conditions for Uniqueness of Equilibrium in Networks," *Mathematics of Operations Research*, vol. 30, no. 1, pp. 225–244, 2005.
[19] M. J. Beckmann, C. B. McGuire, and C. B. Winsten, "Studies in the Economics of Transportation," 1955.
[20] R. Meir and D. Parkes, "Playing the Wrong Game: Smoothness Bounds for Congestion Games with Behavioral Biases," *Performance Evaluation Review*, vol. 43, no. 3, pp. 67–70, 2015.

Appendix: Supporting Lemmas

The first lemma is borrowed from [20, Table 2], giving the price of anarchy for homogeneous populations.

Lemma 5.1 (Meir and Parkes, 2015 [20]): Let $\mathcal{G}^p$ be the class of parallel-network affine-cost congestion games. Let every agent $x \in [0, r]$ have sensitivity $s_x = \beta \geq 0$, and let T^{mc} denote the marginal-cost taxation mechanism which assigns taxation functions to every edge e of $\tau_e(f_e) = a_e f_e$. Write the total latency for population with homogeneous sensitivity β as $\mathcal{L}^{\mathrm{nf}}(G, \beta, T^{\mathrm{mc}})$. If $\beta \in [0, 1]$, then the price of anarchy of $\mathcal{G}^p$ for this population is

$$\sup_{G \in \mathcal{G}^p} \frac{\mathcal{L}^{\mathrm{nf}}(G, \beta, T^{\mathrm{mc}})}{\mathcal{L}^*(G)} = \frac{4}{4(1+\beta) - (1+\beta)^2}. \qquad (28)$$

If $\beta \geq 1$, then

$$\sup_{G \in \mathcal{G}^p} \frac{\mathcal{L}^{\mathrm{nf}}(G, \beta, T^{\mathrm{mc}})}{\mathcal{L}^*(G)} = \frac{(1+\beta)^2}{4\beta}. \qquad (29)$$

The second main supporting lemma shows that under the influence of aggressive tolls T^{AGG}, the price of anarchy *on any network* for heterogeneous populations is realized by a homogeneous population of sensitivity S_{U}. This result is unusual in routing games in that it applies to each instance of a routing game, and thus is considerably stronger than many price of anarchy statements (which are typically considered in worst case over a class of routing games).

Lemma 5.2: Let $G \in \mathcal{G}^p$ be a parallel network with affine cost functions. Let s^{U} denote a homogeneous population with sensitivity S_{U}. For any heterogeneous population $s \in \mathcal{S}$,

$$\mathcal{L}^{\mathrm{nf}}(G, s, T^{\mathrm{AGG}}) \leq \mathcal{L}^{\mathrm{nf}}(G, s^{\mathrm{U}}, T^{\mathrm{AGG}}) \qquad (30)$$

Proof: Due to space limitations, we present here only an outline of the proof of Lemma 5.2. Let f^{nf} denote a Nash flow for population s under the influence of T^{AGG}; suppose that f^{nf} has positive flow on exactly $k \leq n$ links, leaving $n - k$ links with 0 flow. Let G_k denote a network with only the k links that have positive traffic in f^{nf}. Clearly, f^{nf} is a Nash flow for G_k as well, so

$$\mathcal{L}^{\mathrm{nf}}(G, s, T^{\mathrm{AGG}}) = \mathcal{L}^{\mathrm{nf}}(G_k, s, T^{\mathrm{AGG}}). \qquad (31)$$

Using techniques from the proof of [2, Lemma 1.2], it can be shown that

$$\mathcal{L}^{\mathrm{nf}}(G_k, s, T^{\mathrm{AGG}}) \leq \mathcal{L}^{\mathrm{nf}}(G_k, s^{\mathrm{U}}, T^{\mathrm{AGG}}). \qquad (32)$$

Finally, the total latencies of homogeneous s^{U} flows are compared between G and G_k, and it can be shown that

$$\mathcal{L}^{\mathrm{nf}}(G_k, s^{\mathrm{U}}, T^{\mathrm{AGG}}) \leq \mathcal{L}^{\mathrm{nf}}(G, s^{\mathrm{U}}, T^{\mathrm{AGG}}). \qquad (33)$$

Combining (31)-(33) completes the proof. ∎