Social Coordination in Unknown Price-Sensitive Populations

Philip N. Brown[1] and Jason R. Marden[2]

Abstract— In this paper, we investigate the relationship between uncertainty and a designer's ability to influence social behavior. Pigovian taxes are a common approach to social coordination. However, guaranteeing efficient behavior typically requires that the system designer has complete knowledge of the user population's sensitivity to taxation. In this paper, we explore the effect of relaxing this requirement in the context of congestion games with affine costs. Focusing on the class of scaled Pigovian taxes, we derive the optimal tolling scheme that minimizes the worst-case efficiency loss under uncertainty in user sensitivity. Furthermore, we derive explicit bounds which highlight how the level of uncertainty in sensitivity degrades performance.

I. INTRODUCTION

As engineered systems and their users become increasingly interconnected, engineers must consider not only system design, but also efficient system utilization [1]. It is widely known that uninfluenced social systems often exhibit suboptimal behavior. This is a well-researched topic in many fields, including electric power markets [2], communication spectrum sharing [3], and traffic congestion [4]. Accordingly, a central design challenge is to influence social behavior through admissible mechanisms to promote efficiency.

To analyze the role of social influence in engineered systems, we model the behavior of a system's users as a game between the users, in which each user chooses his actions to maximize his personal benefit. Equilibrium behavior in which every user acts in this way is known as a Nash equilibrium. It is widely known that Nash equilibria need not be efficient in any global sense. It is common to quantify the efficiency losses inherent in Nash equilibria by a metric known as the "price of anarchy" of a system, defined as the worst-possible ratio between the efficiency of Nash equilibrium behavior and that of optimal behavior [5], [6]. The price of anarchy has been analyzed for a variety of engineered systems [7], [8]. For example, the price of anarchy for congestion on certain road networks has been shown to be $4/3$ [9]. This leads to the central question: how may we influence social behavior to reduce the price of anarchy?

A common methodology for increasing the efficiency of social systems is known as a Pigovian tax [10]. Under a Pigovian tax, users are charged a tax equal to the negative

effect they have on other users; thus, each user is incentivized to choose actions that maximize the public benefit [11]. Generally, Pigovian taxes are attractive since they guarantee that Nash equilibrium behavior exactly equals system optimal behavior. This type of taxation strategy has been suggested as a solution to industrial pollution [12], harmful speculative trading in financial markets [13], and road congestion [14], to name a few.

We turn to the problem of road congestion to illustrate a setting in which Pigovian taxes apply. In routing games, an equilibrium is known as a Nash flow; it is well-known that Nash flows are often inefficient [15]. A central research focus has been to charge tolls that influence drivers to make more efficient routing choices. In the congestion setting, a Pigovian tax is known as a marginal-cost toll; each driver is charged a toll equivalent to the congestion caused by an additional unit of traffic [16], [17]. This is a fully-distributed tolling system: a road's toll depends only on the road's congestion properties.

An inherent challenge is that congestion is experienced in units of time, but tolls are charged in units of money. If the designer lacks perfect information regarding drivers' price-sensitivity, it may not be possible to levy an accurate toll. For example, if driver price-sensitivities are unknown, a Nash flow may be as much as 33% more congested than an optimal flow [9]. On the other hand, if the price-sensitivity of all drivers is known perfectly, Pigovian tolls cause all Nash flows to be optimal [17], [18].

Recent research has sidestepped this price-sensitivity issue by fixing all system variables and deriving a tolling strategy that optimizes a fixed instance of the system [19]. It has been shown that under a general class of congestion games, this type of fixed tolling can optimize Nash equilibria [20]. While these results represent successes, they rely strongly on a complete characterization of the network structure, user capabilities, and price sensitivities. It is unknown how variations in these parameters degrade a designer's ability to mitigate efficiency losses through tolls. Furthermore, fixed tolls typically rely on global information and do not provide a distributed solution.

With these limitations in mind, we initiate a study on the relationship between knowledge of price sensitivity and a designer's ability to influence social behavior with tolls. Specifically, we investigate the set of network routing games such that each edge in a network possesses a congestion function that is affine in the number of edge users. We analyze worst-case Nash flow efficiency in these games over a range of user price-sensitivities. Affine-cost congestion games with known sensitivities are common in the literature [18], [9],

[15]; however, to our knowledge, this paper is the first price sensitivity analysis of its kind for congestion games.

For this class of games, under mild assumptions, we show in Theorem 3.2 that the unique marginal-cost-toll scale factor that minimizes worst-case efficiency losses is the geometric mean of the bounds of the user population's price-sensitivity. This is a distributed tolling rule: a road's toll depends only on its congestion properties and the amount of traffic on the road; it is independent of the overall network topology.

Furthermore, we show that this tolling rule always increases efficiency with respect to the un-tolled network, and we provide a tight bound on the price of anarchy for games using our tolling methodology. This price of anarchy depends only on the ratio of the user population's sensitivity bounds; it is independent of cost functions, network topology, and overall traffic rate. For example, if the price sensitivity is only known within an order of magnitude, our result gives a price of anarchy of 1.083. This is a marked improvement from the the known price of anarchy of 4/3 for un-tolled affine-cost congestion games [9].

II. MODEL

A. Network Definition

Consider a network (V, E), which consists of a vertex set V and edge set $E \subseteq (V \times V)$. Each edge $e \in E$ is associated with a latency function of the affine form

$$l_e(f_e) = d_e f_e + c_e, \tag{1}$$

where $f_e \geq 0$ designates the flow on edge e, and $d_e \geq 0$ and $c_e \geq 0$ are edge-specific constants.

In this paper, we focus on the efficiency of routing induced by individuals' selfish behavior. Specifically, we focus on the setting in which there is a population of users seeking to traverse the network from a common source to a common destination. Each user chooses his path from a common set of paths $\mathcal{P} = \{p_1, p_2, \ldots, p_n\}$ where each path $p \in \mathcal{P}$ consists of a set of edges, i.e., $\mathcal{P} \subseteq 2^E$. We define the total mass of the population to be $r > 0$. Using these definitions, let a *game* G be defined by the tuple $G = (V, E, \mathcal{P}, \{l_e\}_{e \in E}, r)$, and let $\mathcal{G}$ represent the class of all such games G.

B. Network Flows

For a game G, we call a flow $f = (f_{p_1}, f_{p_2}, \ldots, f_{p_n})$ *feasible* if $\forall p \in \mathcal{P}$, $f_p \geq 0$ and $\sum_{p \in \mathcal{P}} f_p = r$, where f_p denotes the flow on path p. We denote the set of feasible flows for game G as $\mathcal{F}(G)$. For a given flow f, the latency $l_p(f)$ of any path $p \in \mathcal{P}$ is the sum of the latencies of edges in that path:

$$l_p(f) = \sum_{e \in p} l_e(f_e), \tag{2}$$

where f_e is the flow on edge e given f, i.e., $f_e = \sum_{p \in \mathcal{P}: e \in p} f_p$. Thus, the total latency on the network is the flow-weighted sum of all path latencies:

$$L(f) = \sum_{p \in \mathcal{P}} f_p \cdot l_p(f_p). \tag{3}$$

Users are concerned with minimizing their latencies, so the cost J_p to a user of any path $p \in \mathcal{P}$ is the path latency:

$$J_p(f) = \sum_{e \in p} l_e(f_e). \tag{4}$$

We focus in particular on the setting in which each user chooses the path with the lowest cost given the choices of others. A routing obtained in this way is known as a Nash flow [15], which corresponds to a feasible flow $f^{\mathrm{ne}} \in \mathcal{F}(G)$ such that for all $p_i, p_j \in \mathcal{P}$

$$f_{p_i}^{\mathrm{ne}} > 0, f_{p_j}^{\mathrm{ne}} > 0 \implies J_{p_i}(f^{\mathrm{ne}}) = J_{p_j}(f^{\mathrm{ne}}), \tag{5}$$

$$f_{p_i}^{\mathrm{ne}} > 0, f_{p_j}^{\mathrm{ne}} = 0 \implies J_{p_i}(f^{\mathrm{ne}}) \leq J_{p_j}(f^{\mathrm{ne}}). \tag{6}$$

We define an *optimal* flow on G as

$$f^{\mathrm{opt}} = \inf_{f \in \mathcal{F}(G)} L(f).$$

Then we define the price of anarchy of $\mathcal{G}$ as

$$\mathrm{PoA}(\mathcal{G}) = \sup_{G \in \mathcal{G}} \frac{L(f^{\mathrm{ne}} : G)}{L(f^{\mathrm{opt}} : G)}, \tag{7}$$

Where $(f^{\mathrm{ne}} : G)$ and $(f^{\mathrm{opt}} : G)$ denote Nash and optimal flows for G, respectively.

C. Tolling for Known Homogeneous Populations

The network owner charges a toll $\tau_e(f_e)$ on each link in the network in an attempt to influence drivers to make efficient routing choices, leading to a modified path cost of

$$J_p^\tau(f) = \sum_{e \in p} \left[l_e(f_e) + \tau_e(f_e) \right]. \tag{8}$$

A marginal-cost toll for latency functions of the form in (1) is of the form

$$\tau_e(f_e) = d_e f_e. \tag{9}$$

In [15], it is shown that tolls of the form in (9) always induce optimal behavior, doing so in a distributed manner: each edge toll is a function only of the flow and congestion properties of that edge.

D. Tolling for Unknown Homogeneous Populations

In real systems, it is often true that users exhibit price-sensitivities that are unknown or that vary from day to day. We model this price sensitivity with a parameter $\beta \geq 0$. Intuitively, high β implies high aversion to toll. Incorporating price sensitivity, we represent the cost of a path generally as

$$J_p^\tau(f) = \sum_{e \in p} \left[l_e(f_e) + \beta \tau_e(f_e) \right]. \tag{10}$$

In the case when β is known precisely, an optimal marginal-cost toll is of the form $\tau_e(f_e) = d_e f_e / \beta$, since then (10) resolves to (8). However, suppose the sensitivity is some unknown $\beta \in [\beta_{\mathrm{L}}, \beta_{\mathrm{U}}]$. We desire to exploit the distributed nature of marginal-cost tolls, so we investigate tolling functions of the form

$$\tau_e(f_e) = \mu d_e f_e, \tag{11}$$

where our control parameter $\mu \geq 0$ is a common factor for every edge that scales with our estimate of the true β of the

population. We call tolls of this form *scaled marginal-cost tolls*. These tolls further modify (4) to give the final form of our path cost function

$$J_p^\tau(f) = \sum_{e\in p} [(1 + \beta\mu)d_e f_e + c_e].$$ (12)

These modified cost functions combined with (5) and (6) induce a new Nash flow $f^{\rm ne}$. For simplicity of notation, we denote the total latency of a flow $f^{\rm ne}$ given μ, G, and β as $L^{\rm ne}(\mu; G, \beta)$.

Therefore, given a game G, we seek to characterize the scaled marginal-cost tolls which minimize the worst-case total latency for G given a range of user price sensitivities. That is, we seek to characterize solutions μ^* to the optimization problem

$$L^{\rm ne}(\mu^*; G, \beta) = \inf_{\mu\geq 0}\left(\sup_{\beta\in[\beta_{\rm L}, \beta_{\rm U}]} (L^{\rm ne}(\mu; G, \beta))\right).$$ (13)

III. MAIN RESULT

We first make the following assumption on the utilization of the network.

Definition 3.1: A game $G = (V, E, \mathcal{P}, \{l_e\}_{e\in E}, \bar{r})$ is *fully-utilized* if $\forall r \geq \bar{r}$, $\forall p_i \in \mathcal{P}$, $f_{p_i}^{\rm ne}(\mu = 0; G, \beta) > 0$. That is, all paths are in use in an un-tolled Nash flow, and increasing r will not cause any path flow to decrease to zero. Simple arguments can show that this definition applies intuitively to a large class of networks, including all networks of parallel links. In Section V-A we discuss some implications of this definition.

Theorem 3.2: Consider any game G that satisfies Definition 3.1. Let

$$\mu^* = \frac{1}{\sqrt{\beta_{\rm L}\beta_{\rm U}}}.$$ (14)

Then μ^* uniquely satisfies

$$L^{\rm ne}(\mu^*; G, \beta) = \inf_{\mu\geq 0}\left(\sup_{\beta\in[\beta_{\rm L}, \beta_{\rm U}]} (L^{\rm ne}(\mu; G, \beta))\right),$$

and $L(\mu^*; G, \beta) < L(0; G, \beta)$. Furthermore, let $\mathcal{G}^*$ represent the class of all games G satisfying Definition 3.1. For $\beta \in [\beta_{\rm L}, \beta_{\rm U}]$,

$$\text{PoA}(\mu^*; \mathcal{G}^*, \beta) = \frac{4\left(1 + \sqrt{\beta_{\rm L}/\beta_{\rm U}} + \beta_{\rm L}/\beta_{\rm U}\right)}{3\left(1 + \sqrt{\beta_{\rm L}/\beta_{\rm U}}\right)^2}.$$ (15)

Thus, μ^* ensures that the worst-case total latency of a game is minimized for any range-restricted β. Note that μ^* depends only on $\beta_{\rm L}$ and $\beta_{\rm U}$; it is independent of G. This allows us to define a distributed rule for our edge tolls so that we can set each edge's toll without knowledge of the network topology or total traffic rate, and guarantees a reduction in total latency from the un-tolled network.

Our expression for the price of anarchy in (15) succinctly characterizes the relationship between efficiency and uncertainty with respect to price sensitivity. Figure 1 shows how the price of anarchy varies with $\beta_{\rm L}/\beta_{\rm U}$. Note that when $\beta_{\rm L}/\beta_{\rm U} = 0$, meaning that it may not be possible to influence
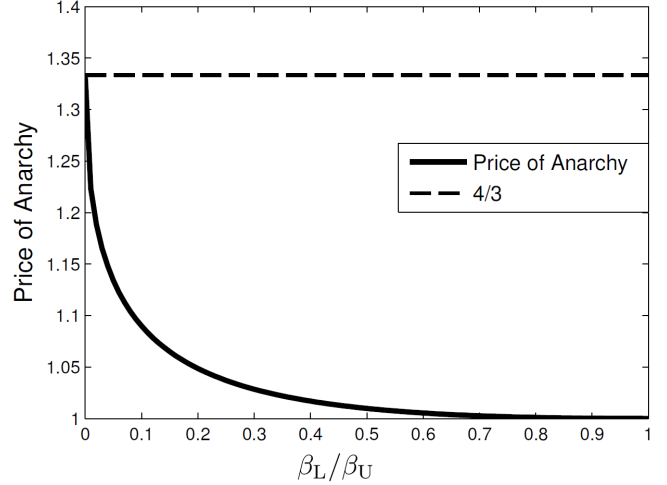

Fig. 1. Price of anarchy for games using our scaled marginal-cost tolling rule, as a function of the precision with which we know the user population's price-sensitivity.

behavior with tolls, (15) resolves to the well-known price of anarchy of $4/3$ derived in [9]. On the other hand, if $\beta_{\rm L} = \beta_{\rm U} > 0$, meaning that the price-sensitivity of all users is known precisely, we obtain a price of anarchy of 1.

Though imperfect information does degrade efficiency, the figure shows that high levels of efficiency are maintained even for a relatively high level of uncertainty. For example, even when β is only known within an order of magnitude, the price of anarchy is 1.083.

IV. PROOF OF RESULTS

The proof proceeds as follows:

1) In Lemma 4.1 we show that in games satisfying Definition 3.1, the flow on any path can be decomposed as the sum of a term depending on r and a term depending on μ.
2) In Lemma 4.2 we show that for such games, Lemma 4.1 implies that scaled marginal-cost tolls always decrease latency by a quantity that is independent of the traffic rate r.
3) Lastly, we show that the properties proved in Lemma 4.2 lead to a characterization of the unique marginal-cost toll scale factor that minimizes worst-case total latency for uncertain sensitivity. We then derive the price of anarchy under our optimal marginal-cost tolls.

Our proof utilizes the following definition: for any $p_i, p_j \in \mathcal{P}$,

$$d_{ij} = \begin{cases} \sum_{e\in p_i\cap p_j} d_e, & p_i \cap p_j \neq \emptyset \\ 0, & p_i \cap p_j = \emptyset \end{cases}.$$ (16)

Intuitively, d_{ij} is the amount that $f_{p_i}^{\rm ne}$ affects $J_{p_j}^\tau(f^{\rm ne})$. Note that $d_{ij} = d_{ji}$.

Lemma 4.1: For any game G satisfying Definition 3.1, for any $\mu \geq 0$, the Nash flow on each path $p_i \in \mathcal{P}$ is of the form

$$f_{p_i}^{\rm ne} = rR_i + \frac{1}{1 + \mu\beta}M_i,$$ (17)

where the coefficients $R_i \in \mathbb{R}$ and $M_i \in \mathbb{R}$ have no dependence on μ, r, or β, and satisfy the following equations:

$$\sum_{p_i \in \mathcal{P}} R_i = 1, \tag{18}$$

$$\sum_{p_i \in \mathcal{P}} M_i = 0, \tag{19}$$

$$\sum_{p_k \in \mathcal{P}} R_k d_{ik} = \sum_{p_k \in \mathcal{P}} R_k d_{jk}, \tag{20}$$

$$\sum_{p_k \in \mathcal{P}} M_k d_{ik} + \sum_{e \in p_i} c_e = \sum_{p_k \in \mathcal{P}} M_k d_{jk} + \sum_{e \in p_j} c_e, \tag{21}$$

where $p_i, p_j, p_k \in \mathcal{P}$.

Proof: First, we prove the decomposition of the path flows. Then we show that due to the properties induced by Definition 3.1, the coefficients R_i and M_i satisfy (18)-(21).

Combining (12) and (II-B), the cost of path p_i is

$$J_{p_i}^\tau(f^{\mathrm{ne}}) = \sum_{e \in p_i} \left((1 + \beta\mu) d_e \left(\sum_{p_k : e \in p_k} f_{p_k}^{\mathrm{ne}} \right) + c_e \right). \tag{22}$$

Define $\bar{d}_e(p)$ in the following way:

$$\bar{d}_e(p) = \begin{cases} d_e, & e \in p \\ 0, & e \notin p. \end{cases}$$

Now, (22) can be written as

$$\begin{aligned} J_{p_i}^\tau(f^{\mathrm{ne}}) &= (1 + \beta\mu) \sum_{e \in p_i} \sum_{p_k \in \mathcal{P}} \bar{d}_e(p_k) f_{p_k}^{\mathrm{ne}} + \sum_{p_i \in \mathcal{P}} c_e \\ &= (1 + \beta\mu) \sum_{p_k \in \mathcal{P}} f_{p_k}^{\mathrm{ne}} \sum_{e \in p_i} \bar{d}_e(p_k) + \sum_{p_i \in \mathcal{P}} c_e \\ &= (1 + \beta\mu) \sum_{p_k \in \mathcal{P}} f_{p_k}^{\mathrm{ne}} d_{ik} + \sum_{e \in p_i} c_e. \end{aligned} \tag{23}$$

Since f^{ne} is a Nash flow and every path flow is positive, by (5) all path costs are equal, so $\forall p_i, p_j \in \mathcal{P}$,

$$\sum_{p_k \in \mathcal{P}} f_{p_k}^{\mathrm{ne}} (d_{ik} - d_{jk}) = \frac{1}{1 + \beta\mu} \left[\sum_{e \in p_j} c_e - \sum_{e \in p_i} c_e \right]. \tag{24}$$

Thus, each $\{i, j\}$ defines an equation of the form $\sum_{p_k \in \mathcal{P}} f_{p_k}^{\mathrm{ne}} a_{ij} = b_{ij}$. We take any distinct[1] $|\mathcal{P}| - 1$ of these equations and combine them with $\sum_{p_k \in \mathcal{P}} f_{p_k}^{\mathrm{ne}} = r$ to obtain a $|\mathcal{P}|$-dimensional system of linear equations, which can be described by the following matrix equation:[2]

$$A f^{\mathrm{ne}} = b. \tag{25}$$

A solution to (25) represents a Nash flow for (G, β, μ), and in [21] it is shown Nash flows always exist for congestion games of the type considered in this paper. Since b is of the form

$$b = \begin{bmatrix} b_1/(1 + \beta\mu) \\ \vdots \\ b_{|\mathcal{P}|-1}/(1 + \beta\mu) \\ r \end{bmatrix}, \tag{26}$$

any solution f^{ne} must be a linear combination[3] of $1/(1+\beta\mu)$ and r. That is, $\forall p_i \in \mathcal{P}$, $f_{p_i}^{\mathrm{ne}}$ can be represented as shown in (17). Note that $\{R_i\}$ and $\{M_i\}$ have no dependence on μ, r, or β.

Clearly, since G satisfies Definition 3.1, R_i must be nonnegative for every path $p_i \in \mathcal{P}$. Note also that $\forall \mu > 0$, $f_{p_i}^{\mathrm{ne}}(\mu; G, \beta) > 0$ since $f_{p_i}^{\mathrm{ne}}(0; G, \beta) > 0$. That is, (17) is robust to increases in tolls.

Next, we derive two important properties of R_i and M_i. Substituting (17) into $\sum_{p_i \in \mathcal{P}} f_{p_i}^{\mathrm{ne}} = r$, we obtain

$$r \left(\sum_{p_i \in \mathcal{P}} R_i - 1 \right) + \frac{1}{1 + \mu\beta} \left(\sum_{p_i \in \mathcal{P}} M_i \right) = 0. \tag{27}$$

Since (27) must hold for any r and for all $\mu \geq 0$, we obtain the proof of (18) and (19).

Now, substituting (17) into (24) and simplifying, we obtain

$$r \sum_{p_k \in \mathcal{P}} R_k (d_{ik} - d_{jk})$$
$$+ \frac{1}{1 + \mu\beta} \left[\sum_{p_k \in \mathcal{P}} M_k (d_{ik} - d_{jk}) + \sum_{e \in p_j} c_e - \sum_{e \in p_i} c_e \right] = 0. \tag{28}$$

Similarly to above, since (28) must hold for any arbitrarily large r and for all $\mu \geq 0$, we obtain the proof of (20) and (21), which completes the proof. ∎

Next, we show that rate and tolls have independent effects on the total latency. This will allow us to define tolling rules without knowledge of r.

Lemma 4.2: For any game G satisfying Definition 3.1, there exists a function $L_r(r)$ depending only on r, and a constant $K_\mu \geq 0$, such that

$$L^{\mathrm{ne}}(\mu; G, \beta) = L_r(r) - \frac{\mu\beta}{(1 + \mu\beta)^2} K_\mu. \tag{29}$$

Proof: First, we show the decomposition of the total latency of a Nash flow.

Substitute (17) into (3) and simplify to obtain

$$L^{\mathrm{ne}}(\mu; G, \beta) = L_r(r) + \frac{r}{1 + \mu\beta} \zeta + \eta, \tag{30}$$

where

$$L_r(r) = \sum_{p_i \in \mathcal{P}} \left[r^2 \sum_{p_j \in \mathcal{P}} d_{ij} R_i R_j + r \sum_{e \in p_i} c_e R_i \right], \tag{31}$$

$$\zeta = \sum_{p_i \in \mathcal{P}} M_i \sum_{p_j \in \mathcal{P}} R_j d_{ij} + \sum_{p_j \in \mathcal{P}} M_j \sum_{p_i \in \mathcal{P}} R_i d_{ij}. \tag{32}$$

$$\eta = \sum_{p_i \in \mathcal{P}} \left[\sum_{p_j \in \mathcal{P}} \frac{d_{ij}}{(1 + \mu\beta)^2} M_i M_j + \sum_{e \in p_i} c_e \frac{1}{1 + \mu\beta} M_i \right]. \tag{33}$$

---

[1]Note that $\{i, j\}$ is *not* distinct from $\{j, i\}$.
[2]With abuse of notation, here we represent the tuple of path flows as a column vector.

[3]To see this, perform Gaussian elimination on the augmented matrix $[A\ b]$ to solve for f. Note that matrix row operations only involve sums of elements of b, never products.

Note that ζ depends on both r and μ. Here we show that $\zeta = 0$. Choose any $p_k \in \mathcal{P}$. By Lemma 4.1, $\forall p_i, p_k \in \mathcal{P}$, $\sum_{p_j \in \mathcal{P}} R_j d_{ij} = \sum_{p_j \in \mathcal{P}} R_j d_{kj}$, and analogously, $\forall p_j, p_k \in \mathcal{P}$, $\sum_{p_i \in \mathcal{P}} R_i d_{ij} = \sum_{p_i \in \mathcal{P}} R_i d_{ik}$. Thus, (32) can be written

$$\zeta = \left(\sum_{p_i \in \mathcal{P}} M_i \right) \sum_{p_j \in \mathcal{P}} R_j d_{kj} + \left(\sum_{p_j \in \mathcal{P}} M_j \right) \sum_{p_i \in \mathcal{P}} R_i d_{ik}.$$

By Lemma 4.1, we know that each term in parentheses equals 0, so $\zeta = 0$. Thus, the total latency is the sum of a function dependent on r and a function dependent on μ and β, thus decoupling the effects of rate and tolls.

Consider now the $1/(1 + \mu\beta)^2$ term of (33) and denote it K_μ:

$$K_\mu = \sum_{p_i \in \mathcal{P}} M_i \sum_{p_j \in \mathcal{P}} d_{ij} M_j. \qquad (34)$$

By (21) in Lemma 4.1, $\forall p_i, p_k \in \mathcal{P}$, we know that

$$\sum_{p_j \in \mathcal{P}} d_{ij} M_j = \sum_{p_j \in \mathcal{P}} d_{kj} M_j + \sum_{e \in p_k} c_e - \sum_{e \in p_i} c_e.$$

Accordingly, we can write K_μ as

$$\left(\sum_{p_i \in \mathcal{P}} M_i \right) \left[\sum_{p_j \in \mathcal{P}} d_{jk} M_j + \sum_{e \in p_k} c_e \right] - \sum_{p_i \in \mathcal{P}} \sum_{e \in p_i} c_e M_i.$$

By Lemma 4.1 we know that $\sum_{p_i \in \mathcal{P}} M_i = 0$. Therefore K_μ is simply

$$K_\mu = - \sum_{p_i \in \mathcal{P}} \sum_{e \in i} c_e M_i.$$

Factoring out K_μ, we can then express (30) as

$$L^{\mathrm{ne}}(\mu; G, \beta) = L_r(r) - \frac{\mu\beta}{(1 + \mu\beta)^2} K_\mu. \qquad (35)$$

To complete the proof of the lemma, we show that $K_\mu \geq 0$. Equation (35) clearly shows that

$$L^{\mathrm{ne}}(0; G, \beta) = L_r(r). \qquad (36)$$

That is, $L_r(r)$ represents the total latency of the un-tolled system. Equations (35) and (36) show that for a given G, if some toll decreases latency, then all tolls must decrease latency. Since $\mu = 1/\beta$ always results in optimal routing[4] (i.e., a reduction in latency), it must be true that $K_\mu \geq 0$. Thus, scaled marginal-cost tolls always decrease total latency. ∎

Completion of Proof of Theorem 3.2. Consider the partial derivative of (29) with respect to β:

$$\frac{\partial L^{\mathrm{ne}}(\mu; G, \beta)}{\partial \beta} = \left(\frac{\mu(\mu\beta - 1)}{(1 + \mu\beta)^3} \right) K_\mu. \qquad (37)$$

From (37), it is clear that $L^{\mathrm{ne}}(\mu; G, \beta)$ has a minimum at $\beta = 1/\mu$, so the end-points of the β-range give the worst possible latencies:

$$\sup_{\beta \in [\beta_{\mathrm{L}}, \beta_{\mathrm{U}}]} \{ L^{\mathrm{ne}}(\mu; G, \beta) \} = \max \{ L^{\mathrm{ne}}(\mu; G, \beta_{\mathrm{L}}), L^{\mathrm{ne}}(\mu; G, \beta_{\mathrm{U}}) \}.$$

[4]See Section II-D.

Consider now the partial derivative of $L^{\mathrm{ne}}(\mu; G, \beta)$ with respect to μ:

$$\frac{\partial L^{\mathrm{ne}}(\mu; G, \beta)}{\partial \mu} = \left(\frac{\beta(\mu\beta - 1)}{(1 + \mu\beta)^3} \right) K_\mu. \qquad (38)$$

Clearly, for any $\beta > 0$, $L^{\mathrm{ne}}(\mu; G, \beta)$ has a minimum at $\mu = 1/\beta$. Furthermore, since $L^{\mathrm{ne}}(1/\beta; G, \beta) = L(f^{\mathrm{opt}})$, it must be true that $L^{\mathrm{ne}}(1/\beta_{\mathrm{L}}; G, \beta_{\mathrm{L}}) = L^{\mathrm{ne}}(1/\beta_{\mathrm{U}}; G, \beta_{\mathrm{U}})$. Thus, since $L^{\mathrm{ne}}(\mu; G, \beta)$ is continuous in μ, there must exist some $\mu^* \in [1/\beta_{\mathrm{U}}, 1/\beta_{\mathrm{L}}]$ such that $L^{\mathrm{ne}}(\mu^*; G, \beta_{\mathrm{L}}) = L^{\mathrm{ne}}(\mu^*; G, \beta_{\mathrm{U}})$. It can easily be verified from (29) that μ^* is as defined in (14) in the theorem statement.

Now we bound the inefficiency of a game G under tolls as defined in (14). Since we know that an un-tolled latency can never be more than $4/3$ times an optimal latency, from (29) we can write

$$\frac{L^{\mathrm{ne}}(0; G, \beta)}{L(f^{\mathrm{opt}})} = \frac{L_r(r)}{L_r(r) - \frac{1}{4} K_\mu} \leq \frac{4}{3}. \qquad (39)$$

This allows us to compute a bound on K_μ:

$$K_\mu \leq L_r(r). \qquad (40)$$

It follows algebraically that for μ^* as defined in (14), $\beta \in [\beta_{\mathrm{L}}, \beta_{\mathrm{U}}]$, and $\mathcal{G}^*$ as defined in the statement of the theorem, the Price of Anarchy is given by the expression in (15). We show that this bound is tight in Section V-B with Pigou's Example. ∎

V. DISCUSSION

A. A Note on Full Utilization

It remains an open question as to the importance of Definition 3.1 for the above results. It is straightforward to show that many networks intuitively satisfy the definition, such as all networks that are composed entirely of simple combinations of parallel and series subnetworks. However, there do exist networks for which Definition 3.1 never applies, notably the network used in Braess's Paradox [22]. This does not imply that it is impossible to achieve efficiency gains for Braess's network with a distributed tolling rule; it simply means that the tolling rule specified in this paper is not guaranteed to be the solution to the optimization problem (13). Definition 3.1 is a sufficient condition for Theorem 3.2, but it should not be interpreted as necessary.

B. Pigou's Example

Consider a network in which drivers have access to two routes: the first, a congestion-sensitive route; the second, a costly constant-latency route; this is the canonical routing problem known as Pigou's Example [18], illustrated in Figure 2. Let $r = 1$. Without tolls, all traffic[5] prefers route 1 since no driver can do better by using route 2, and the total latency $L^{\mathrm{ne}}(0; G, \beta) = 1$.

[5]Note that this example does not strictly satisfy Definition 3.1 since $f_{p_2} = 0$. This will not adversely impact our analysis, because we may consider the case in which $r = 1 + \epsilon$ for $\epsilon > 0$. Then $f_{p_2} = \epsilon$ and $f_{p_1} = 1$, so Definition 3.1 is satisfied. We take the limit as $\epsilon \to 0$, and all analysis proceeds as before.
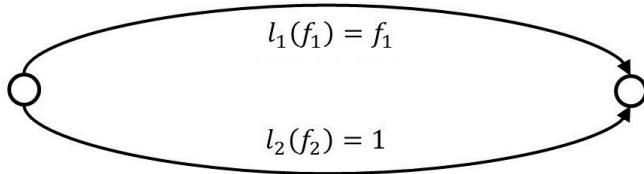
Fig. 2. Pigou's Example. Network used in Section V-B to demonstrate the price of anarchy for scaled marginal-cost tolls and uncertain user populations.

On the other hand, the flow which minimizes the total congestion is $f_{p_1} = f_{p_2} = 0.5$; now half the population experiences a latency of 0.5 and half 1, so the total latency is $L(f^{\mathrm{opt}}) = 3/4$; the price of anarchy for this instance is thus $4/3$.

Assume that all users share some value of $\beta \in [\beta_{\mathrm{L}}, \beta_{\mathrm{U}}]$. Charge a toll on the first link of $\tau_1(f) = f_1\mu$, where $\mu \geq 0$. The Nash flow induced by this toll is such that $f_{p_1} = \mu/(\mu + \beta)$ and $f_{p_2} = \beta/(\mu + \beta)$. Thus the total latency is

$$L^{\mathrm{ne}}(\mu; G, \beta) = \left(\frac{\mu}{\mu + \beta}\right)^2 + \frac{\beta}{(\mu + \beta)} \qquad (41)$$

It is easy to verify that $\mu = (\beta_{\mathrm{L}}\beta_{\mathrm{U}})^{-1/2}$ as prescribed by Theorem 3.2 minimizes the worst-case total latency for any $\beta \in [\beta_{\mathrm{L}}, \beta_{\mathrm{U}}]$ and that this worst-case total latency is

$$L^{\mathrm{ne}}(\mu^*; G, \beta_{\mathrm{L}}) = \frac{\left(1 + \sqrt{\beta_{\mathrm{L}}/\beta_{\mathrm{U}}} + \beta_{\mathrm{L}}/\beta_{\mathrm{U}}\right)}{\left(1 + \sqrt{\beta_{\mathrm{L}}/\beta_{\mathrm{U}}}\right)^2}. \qquad (42)$$

Since $L(f^{\mathrm{opt}}) = 3/4$, (42) proves the tightness of the price of anarchy specified in Theorem 3.2.

VI. CONCLUSIONS

Effective design for social coordination requires information about a system's user population. If this information is inaccurate, an attempt by the designer to influence user behavior may result in users making inefficient choices. It is essential to quantify the effect of these inaccuracies in order to ensure that attempts to mitigate efficiency losses do not inflame the problem.

Toward this end we have demonstrated a positive result: we have developed a distributed tolling strategy for certain road network settings that always reduces Nash flow road congestion, even if the user population's price sensitivity is not precisely known. In our model, uncertainties in price sensitivity did reduce efficiency, but relatively high efficiency was maintained even for high levels of uncertainty.

This area has many open questions. In the congestion setting, it remains to be shown whether these results extend to more general congestion games. For example, future research will focus on eliminating the requirement that games satisfy Definition 3.1. It is not known whether the efficiency guarantees proved in this paper will still apply after the removal of this requirement. Another important issue is that of heterogeneity. Our model strongly assumes homogeneity of the population's price sensitivities; as research progresses,

we hope to relax this assumption. Homogeneity is rare in practice, and incorporating its effects into our models will be an important step.

This style of analysis, in which we optimize our incentive scheme over a range of user price sensitivities, could be applied to a wide variety of problems, both engineering and economic in nature. From electric power markets to energy efficiency tax credits, price sensitivity plays an integral role in many engineering and social systems. This paper represents an initial foray into a broader analysis of the impact of price sensitivity uncertainty on designers' ability to promote efficiency through financial incentives.

REFERENCES

[1] G. E. Bolton and A. Ockenfels, "Behavioral Economic Engineering," *Journal of Economic Psychology*, vol. 33, pp. 665–676, June 2012.

[2] C. Silva, "Application of Mechanism Design to Electric Power Markets (Republished)," *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 862–869, 2001.

[3] D. Niyato and E. Hossain, "Competitive Pricing for Spectrum Sharing in Cognitive Radio Networks: Dynamic Game, Inefficiency of Nash Equilibrium, and Collusion," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 192–202, 2008.

[4] S. Morrison, "A Survey of Road Pricing," *Transportation Research Part A: General*, no. 2, pp. 87–97, 1986.

[5] C. Papadimitriou, "Algorithms, Games, and the Internet," in *Proc. of the 28th International Colloquium on Automata, Languages and Programming*, 2001.

[6] E. Koutsoupias and C. Papadimitriou, "Worst-Case Equilibria," *Computer Science Review*, vol. 3, pp. 65–69, May 2009.

[7] R. Johari and J. N. Tsitsiklis, "Efficiency Loss in a Network Resource Allocation Game," *Mathematics of Operations Research*, vol. 29, pp. 407–435, Aug. 2004.

[8] H. Youn, M. Gastner, and H. Jeong, "Price of Anarchy in Transportation Networks: Efficiency and Optimality Control," *Physical Review Letters*, vol. 101, pp. 128701–1 to 4, Sept. 2008.

[9] T. Roughgarden, "How Bad Is Selfish Routing?," *Journal of the ACM JACM (2002)*, vol. 49, no. 2, pp. 236–259, 2002.

[10] N. G. Mankiw, "Smart Taxes: An Open Invitation to Join the Pigou Club," *Eastern Economic Journal*, vol. 35, no. 1, pp. 14–23, 2009.

[11] W. Sandholm, "Negative Externalities and Evolutionary Implementation," *The Review of Economic Studies*, vol. 72, pp. 885–915, July 2005.

[12] W. Baumol, "On Taxation and the Control of Externalities," *The American Economic Review*, vol. 62, no. 3, pp. 307–322, 1972.

[13] J. Stiglitz, "Using Tax Policy to Curb Speculative Short-Term Trading," *Journal of Financial Services Research*, vol. 3, no. 2-3, pp. 101–115, 1989.

[14] I. W. H. Parry, M. Walls, and W. Harrington, "Automobile Externalities and Policies," *SSRN Electronic Journal*, 2007.

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

[16] R. Cole, Y. Dodis, and T. Roughgarden, "How Much Can Taxes Help Selfish Routing?," *Journal of Computer and System Sciences*, vol. 72, pp. 444–467, May 2006.

[17] W. Sandholm, "Evolutionary Implementation and Congestion Pricing," *The Review of Economic Studies*, vol. 69, no. 3, pp. 667–689, 2002.

[18] A. C. Pigou, *The Economics of Welfare*. 1920.

[19] R. Cole, Y. Dodis, and T. Roughgarden, "Pricing Network Edges for Heterogeneous Selfish Users," *Proc. of the thirty-fifth ACM symp. on Theory of computing - STOC '03*, p. 521, 2003.

[20] L. Fleischer, K. Jain, and M. Mahdian, "Tolls for Heterogeneous Selfish Users in Multicommodity Networks and Generalized Congestion Games," *45th Annu. IEEE Symp. on Foundations of Computer Science*, pp. 277–285, 2004.

[21] 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.

[22] D. Braess, A. Nagurney, and T. Wakolbinger, "On a Paradox of Traffic Planning," *Transportation Science*, vol. 39, pp. 446–450, Nov. 2005.