Incentivizing Efficient Equilibria in Traffic Networks with Mixed Autonomy

Erdem Bıyık*, Daniel A. Lazar*, Ramtin Pedarsani, Dorsa Sadigh

Abstract—Traffic congestion has large economic and social costs. The introduction of autonomous vehicles can potentially reduce this congestion by increasing road capacity via vehicle platooning and by creating an avenue for influencing people's choice of routes. We consider a network of parallel roads with two modes of transportation: (i) human drivers, who will choose the quickest route available to them, and (ii) a ride hailing service, which provides an array of autonomous vehicle route options, each with different prices, to users. We formalize a model of vehicle flow in mixed autonomy and a model of how autonomous service users make choices between routes with different prices and latencies. Developing an algorithm to learn the preferences of the users, we formulate a planning optimization that chooses prices to maximize a social objective. We demonstrate the benefit of the proposed scheme by comparing the results to theoretical benchmarks which we show can be efficiently calculated.

I. INTRODUCTION

Road congestion is a major and growing source of inefficiency, costing drivers in the United States billions of dollars in wasted time and fuel [1]. Autonomous vehicles could improve the efficiency of road usage by forming platoons, thereby increasing road capacity and smoothing traffic flow. The presence of *mixed autonomy*, where human drivers and autonomous vehicles share roads, complicates these improvements – if drivers make their routing decisions *selfishly* and seek to minimize their experienced latency, this can result in suboptimal network performance [2]. Moreover, increasing road capacity by converting human-driven vehicles to autonomous ones can paradoxically worsen average transit user latency [3].

It is therefore important to consider the effect of self-interested users on the system as well as how to *influence* these users to take routes which mitigate this effect. In this paper we consider monetary incentivizes for such prosocial behavior. We consider a setting where a benevolent ride hailing service or social planner sets prices for each route that an autonomous vehicle user may take. The users of the autonomous service will choose their routes, or choose not to travel, based on their time-money tradeoff. The remaining human drivers will choose routes that minimize their latency. The role of the social planner is then to choose prices to optimize some social objective. Our model is an indirect Stackelberg game – a game in which a social planner controls some fraction of the population's actions, where the remainder of the population responds selfishly. However, in our model the planner only controls its portion of the vehicle flow indirectly, via pricing.

To effectively set prices, the social planner needs a model for a) the flow of vehicles on a road, which depends on how many autonomous and human-driven vehicles are on the road, and b) how people make decisions between routes with various prices and latencies. We model the latter based on multinomial logits [4] and model human drivers as selfish agents who reach a *Nash Equilibrium* (also called a Wardrop Equilibrium in the context of transportation networks), a configuration in which no one can decrease their travel time by switching paths [5].

Moreover, we use active preference-based learning [6], via a series of queries, to understand the preferences of autonomous service users. This enables the planner to predict how autonomous service users will react to a set of options with only a relatively low number of training queries. We experimentally verify this method accurately predicts human choice, and our planning algorithm can use this to improve network performance. Furthermore, we provide a theoretical framework which establishes benchmarks for the performance of the algorithm and we show that these benchmarks can be calculated in polynomial time. Finally, we show the efficacy of our pricing scheme in the context of these benchmarks.

Our contributions in this work are as follows:
- We develop a formal mixed autonomy traffic flow model.
- We develop an active preference-based algorithm to learn how different people value time and money in choosing their transportation option. This enables learning a model for people's routing choices in a data-efficient manner.
- We use these models to formulate and solve an optimization for ride hailing service to minimize congestion and maximize the road network utilization while constraining a minimum profit for the service supplier.
- We provide theoretical benchmarks for understanding the performance of the above and prove that we can calculate these benchmarks in polynomial time.
- We validate the learning algorithm and the planning optimization formulation via simulations and user studies. Our results show carefully designed pricing schemes can significantly improve traffic throughput.

Related work. Previous works have shown the potential benefits that autonomous vehicles can have for traffic networks by increasing road capacity through platooning [7], [8], damping shockwaves of slowing vehicles [9], [10], managing merges [11], decongesting freeways in the event of accidents [12], [13], and balancing a supply of vehicles [14]. Relatedly, many works analyze and bound the inefficiency that can arise from network users choosing their routes selfishly [15], including when autonomous vehicles are introduced [3].

As we use financial incentives to influence this behavior,

*Authors contributed equally.

Erdem Bıyık and Dorsa Sadigh are with the Department of Electrical Engineering, Stanford University. Dorsa Sadigh is also with the Department of Computer Science, Stanford University. Daniel Lazar and Ramtin Pedarsani are with the Department of Electrical and Computer Engineering, UC Santa Barbara. `ebiyik@stanford.edu`, `dlazar@ece.ucsb.edu`, `ramtin@ece.ucsb.edu`, `dorsa@stanford.edu`

our formulation is related to work in tolling, some of which consider users with different sensitivities to tolls [16], [17]. [18] considers a congestion game framework and derives tolls which drive users to choose socially optimal strategies for a broad class of user strategy update dynamics. In contrast to many of these works, we consider an empirically validated probabilistic model for human choice [19] which incorporates differences in price sensitivity, and we model vehicle flow on shared roads based on the foundational Fundamental Diagram of Traffic (FDT) [20]. Some works consider heterogeneous and stochastic user utilities (*e.g.* [21], [22]). However, the approaches taken in these works cannot incorporate a FDT-based model for road congestion, in which latency is no longer an increasing function of vehicle flow on a road. Also relatedly, [23] considers a Stackelberg game (meaning a planner can route a portion of the vehicle flow) on parallel roads with flow dictated by the FDT when there is a single vehicle type. Hence, a major novelty of our work lies in that we consider a mixed-autonomy network where autonomous service users have different preferences.

To understand human choice, there has recently been much effort on learning human reward functions which are assumed to be sufficient to model preferences. Inverse reinforcement learning [24]–[26] and preference-based learning [6], [27]–[29] are the most popular choices. In this paper, we employ preference-based learning, a natural fit to our problem. We actively synthesize queries – a non-trivial generalization and extension of [6] – for data-efficiency and better usability.

II. PROBLEM SETTING AND OBJECTIVE

A. Vehicle Flow Model

We assume every road i has a maximum flow. This occurs when traffic is in *free-flow* – when all vehicles travel at the nominal road speed $\bar{v}_i$.

Definition 1. *The* free-flow latency *of a road i, denoted a_i, is the time it takes vehicles to traverse the road in free-flow. With road length denoted d_i, the free-flow latency is $a_i := d_i/\bar{v}_i$.*

Traffic Density vs Traffic Flow: Adding more cars to a road that is already at maximum flow makes the traffic switch from free-flow to a congested regime, which decreases the vehicle flow. In the extreme case, at a certain density $\bar{n}_i$, cars are bumper-to-bumper and vehicle flow stops. The solid lines in Fig. 1(a) – Fundamental Diagram of Traffic [20] – illustrates this phenomenon, where flow increases linearly with respect to density until it hits the critical density. The slope corresponds to the free-flow velocity $\bar{v}_i$ on road i. After the critical point, flow decreases linearly until it is zero at the maximum density. In this figure, $\bar{F}$ is the maximum flow and $\tilde{n}$ denotes the critical density, where the argument is the fraction of vehicles that are autonomous. We will formally define this notation below.
Traffic Flow vs Road Latency: The relationship between vehicle flow and road latency reflects the same free-flow/congested divide. As mentioned above, roads in free-flow have constant latency. In the congested regime, however, latency *increases* as vehicle flow decreases, since a high density of vehicles is required to achieve a low traffic flow. This is represented in Fig. 1(b) with the solid lines.
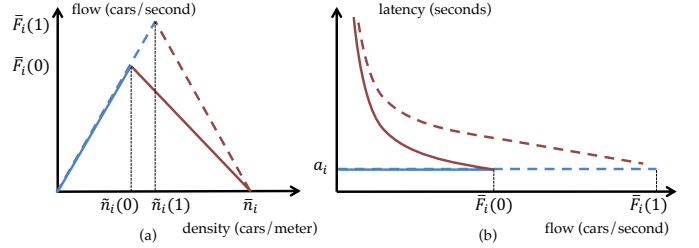

Fig. 1: **(a)** The Fundamental Diagram of Traffic for roads with all human-driven (solid) and all autonomous (dashed) vehicles. In the latter, congestion begins at a higher vehicle density as autonomous vehicles require a shorter headway when following other vehicles. **(b)** The relationship between vehicle flow and latency also changes in the presence of autonomous vehicles. Free-flow speed remains the same but maximum flow on a road increases.

Mixed-Autonomy Roads. We assume that on mixed-autonomy roads, the autonomous vehicles can coordinate with one another and potentially form platoons to help with the efficiency of the road network. We now extend the traffic model above to mixed-autonomy settings as shown with dashed lines in Fig. 1. We define the autonomy level of a road i as the fraction of autonomous vehicle flow on that road: $\alpha_i := \frac{f_i^a}{f_i^a + f_i^h}$ where f_i^a and f_i^h represent the autonomous and human-driven vehicle flow, respectively. Assuming that neither the nominal velocity $\bar{v}_i$ nor the maximum density $\bar{n}_i$ changes, the critical density at which traffic becomes congested will now shift and increase with *autonomy level* α_i, as platooned autonomous vehicles require a shorter headway than human drivers.

To formalize the relationship between autonomy level and critical density on road i, we assume the space occupied by autonomous vehicles and humans at nominal velocity is h_i^a and h_i^h, respectively, with $h_i^a \leq h_i^h$. This inequality reflects the assumption that autonomous vehicles can maintain a short headway, regardless of the type of vehicle they are following. Then, the critical density is

$$\tilde{n}_i(\alpha_i) := \frac{b_i}{\alpha_i h_i^a + (1 - \alpha_i)h_i^h} , \qquad (1)$$

where b_i is the number of lanes on that road. Here the denominator represents the average length from one car's rear bumper to the preceding car's rear bumper when all cars follow the vehicle in front of them with nominal headway. Note that critical density is expressed here as a function of the autonomy level α_i of the road. Since flow increases linearly with density until hitting the critical point, the maximum flow can also be expressed as a function of autonomy level: $\bar{F}_i(\alpha_i) = \bar{v}_i \tilde{n}_i(\alpha_i)$.

The flow on a road, $f_i = f_i^h + f_i^a$ is a function of the density (n_i^h and n_i^a, respectively) of each vehicle type as follows.

$$f_i(n_i^h, n_i^a) :=$$
$$\begin{cases} \bar{v}_i \cdot (n_i^h + n_i^a), & \text{if } n_i^h + n_i^a \leq \tilde{n}_i(\alpha_i) \\ \frac{\bar{v}_i \cdot \tilde{n}_i(\alpha_i) \cdot (\bar{n}_i - (n_i^h + n_i^a))}{\bar{n}_i - \tilde{n}_i(\alpha_i)}, & \text{if } \tilde{n}_i(\alpha_i) \leq n_i^h + n_i^a \leq \bar{n}_i \quad (2) \\ 0, & \text{otherwise} . \end{cases}$$

We can then write the latency as a function of vehicle flow as well as a binary argument s_i, which indicates whether the road is congested [23], [30]:

$$\ell_i(f_i^h, f_i^a, s_i) = \begin{cases} \frac{d_i}{\bar{v}_i} & \text{if } s_i = 0 \\ d_i \left(\frac{\bar{n}_i}{f_i^h + f_i^a} + \frac{\tilde{n}_i(\alpha_i) - \bar{n}_i}{\bar{v}_i \cdot \tilde{n}_i(\alpha_i)} \right) & \text{if } s_i = 1 . \end{cases} \quad (3)$$
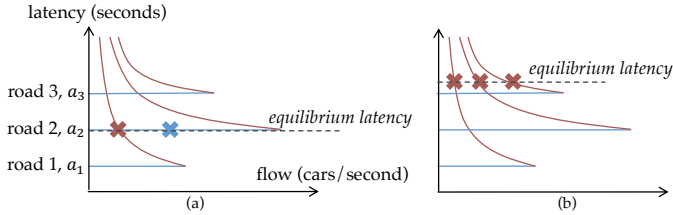
Fig. 2: Some possible equilibria of a three-road network with fixed flow demand. Blue and red lines denote the free-flow and congested regimes, respectively. Equilibria may have (a) one road in free-flow or (b) all roads congested. An equilibrium has an associated *equilibrium latency* experienced by all selfish users. By considering a given equilibrium latency, we can reason about which roads must be congested at that equilibrium as well how much flow is on each road.

Fig. 1(b) illustrates the effect of mixed autonomy on latency.

B. Network Model

Assumption 1. *We consider a network of N parallel roads. We assume that no two roads have the same free-flow latency. We order the indices such that $a_1 < a_2 < \ldots < a_N$.*

The role of the assumption above is explained in the Appendix. We use $[k]$ to denote the set of the first k roads; accordingly, $[N]$ denotes the set of all roads.

We describe the network state by $(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s})$, where $\boldsymbol{f}^{\mathrm{h}}$, $\boldsymbol{f}^{\mathrm{a}} \in \mathbb{R}_{\geq 0}^{N}$ and $\boldsymbol{s} \in \{0, 1\}^{N}$. A *feasible routing* is one for which $f_i^{\mathrm{h}} + f_i^{\mathrm{a}} \leq \bar{F}_i(\alpha_i)$ for all roads, and the flow of each vehicle type on the roads sum to λ^{h} and λ^{a} in the case of inelastic demand, respectively, which denote the total vehicle flow demands. In the case where the demand for autonomous vehicles is elastic, i.e., Section IV, this constraint is relaxed such that the total autonomous flow does not exceed the maximum demand. We are interested in finding a routing, i.e. allocation of vehicles into the roads, that minimizes the total latency experienced by all vehicles while maximizing the total flow of the roads in the case of elastic demand. Further, we constrain this optimization based on *total demand*, *selfishness* or *flexibility* of the vehicles. While selfish drivers always take the quickest road possible, flexible vehicles accept relatively longer latencies. We will formalize these terms in Section III.

C. Human Driver Choice Model

We assume human drivers are selfish, i.e. their only consideration is minimizing their own commute time. This leads to a *Nash Equilibrium* [5], which, on parallel roads, means that if one road has positive flow on it, all other roads must have higher or equal latency. Formally,

$$f_i^{\mathrm{h}} > 0 \implies \ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i) \leq \ell_{i'}(f_{i'}^{\mathrm{h}}, f_{i'}^{\mathrm{a}}, s_{i'}) \; \forall i, i' \in [N] \,.$$

This implies that all selfish users experience the same latency. It is therefore useful to consider the flow-latency diagrams of roads when studying which equilibria exist – by fixing the latency on the y-axis, one can reason about which roads must be congested to achieve that equilibrium. As shown in Fig. 2, equilibria may have one road in free-flow and rest congested, or all may be congested [23], [30].

D. System Objective

In this paper, we are interested in developing a pricing scheme for autonomous vehicles that improves the state of the traffic by alleviating the adverse effects of selfishness. First, we develop two baselines: the first is the case in which all users,

including autonomous vehicles, are selfish and the network reaches a *Nash Equilibrium*. In this case we wish to efficiently calculate the Nash Equilibrium that minimizes overall travel latency. Our goal is to achieve lower travel latencies than this equilibrium with the same amount of flow. In the second baseline, we assume we have limited direct control over the routing of autonomous users – more specifically, we can route autonomous vehicles as we wish as long as the latency they experience is within some range of the quickest route available. A specific case of this baseline with full flexibility of the autonomous vehicles serves as a lower bound to our method.

After developing the two baselines in Section III, we present our pricing method that provides the same benefits as flexible behavior through financial incentives in Section IV. For our pricing scheme, we describe the various facets of the problem. We assume the demand of human drivers is fixed, and the demand of people using the autonomous service is elastic – *if prices and latencies are high, some people may choose not to use the autonomous mobility service.* Our goal is then to simultaneously maximize the number of autonomous service users that can use the road, and minimize the average latency experienced by all the people using the roads.

III. PERFORMANCE BENCHMARKS

Throughout this section, we assume an inelastic demand, i.e., we will route all autonomous and human-driven flow demand, λ^{a} and λ^{h}, into the network. As the demand is inelastic, we are only interested in finding a routing that minimizes the total latency experienced by all vehicles, $C(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) = \sum_{i \in [N]} (f_i^{\mathrm{h}} + f_i^{\mathrm{a}}) \ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i)$, while satisfying the demand, i.e. $\sum_{i \in [N]} f_i^{\mathrm{h}} = \lambda^{\mathrm{h}}$ and $\sum_{i \in [N]} f_i^{\mathrm{a}} = \lambda^{\mathrm{a}}$.

We now make precise the aforementioned notions of selfishness and flexibility. We develop properties of the resulting equilibria, and using those, provide polynomial-time algorithms for computing the benchmark flows.

Selfishness. Human drivers are often thought of as selfish, meaning they will not take a route with long latency if a quicker route is available to them. If all drivers are selfish this leads to a *Nash Equilibrium*, in which no driver can achieve a lower travel time by unilaterally switching routes [5]. This means that all selfish users with the same origin and destination experience the same travel time.

Definition 2. *The **longest equilibrium road** is the road with maximum free-flow latency which has latency equal to the latency experienced by selfish users. Let m_{EQ} denote the index of this road. We use $NE(\lambda^h, \lambda^a, m_{\mathrm{EQ}})$ to denote the set of Nash Equilibria with* longest equilibrium road *having index m_{EQ}.*

Definition 3. *The **longest used road** is the road with maximum free-flow latency that has positive vehicle flow of any type on it. We use m_{ALL} to denote the index of this road; if all vehicles in a network are selfish then $m_{\mathrm{EQ}} = m_{\mathrm{ALL}}$.*

The following lemma will help with the subsequent theoretical results; we defer its proof to the appendix.

Lemma 1. *If the set of Nash Equilibria contains a routing with positive flow only on roads $[m]$, then there exists a routing in*
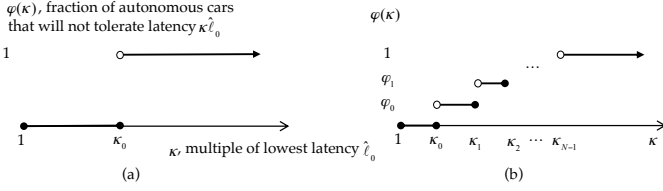
Fig. 3: Flexibility profiles. A fraction $\varphi(\kappa)$ of autonomous users will not accept latency greater than κ times that of the quickest available route. **(a)** Users will tolerate latency of up to κ_0 times that of the quickest route. **(b)** Users have multiple flexibility levels.

the set of Nash Equilibria with positive flow only on roads $[m']$ where $m' \leq m$, and road m' is in free-flow.

We define the set of ***Best-case Nash Equilibria (BNE)*** as the set of feasible routings in equilibrium that minimize the total latency for flow demand $(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}})$, denoted BNE$(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}})$. The following theorem (proof deferred to the appendix) provides properties of the set of BNE for mixed-autonomy roads (for roads with a single vehicle type, see [23]).

Theorem 1. *There exists a road index m_{EQ}^* such that all routings in the set of BNE have the below properties. Further, this index m_{EQ}^* is the minimum index such that a feasible routing can satisfy the properties:*
1) *road m_{EQ}^* is in free-flow,*
2) *roads with index less than m_{EQ}^* are congested with latency $a_{m_{\mathrm{EQ}}^*}$, and*
3) *all roads with index greater than m_{EQ}^* have zero flow.*

As the same latency level can be achieved by varying the autonomy levels of the roads, BNE is not necessarily unique. **Flexibility.** We also wish to find a lower bound for the social cost when some users are willing (or incentivized) to take longer routes. To that end, we use the term *flexibility profile* to refer to the distribution of the degree to which autonomous users are willing to endure longer routes. For computational reasons, we consider flexibility profiles with a finite number of flexibility levels.

Formally, we define $\varphi : \mathbb{R}_{\geq 0} \to [0,1]$ to represent the flexibility profile as a nondecreasing function of a latency value that is mapped to $[0,1]$. A volume of $\varphi(\kappa)\lambda^{\mathrm{a}}$ autonomous flow will reject a route incurring latency κ times the minimum route latency available, which we denote $\ell_{m_{\mathrm{EQ}}}$. If autonomous users have a uniform flexibility level as in Fig. 3(a), we call them κ_0-*flexible* users, where κ_0 is the maximum multiple of the minimum latency that autonomous users will accept. Users may have differing flexibility levels, as in Fig. 3(b). We use K to denote the set of flexibility levels, with cardinality $|K|$. Accordingly, a feasible routing $(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s})$ is in the set of ***Flexible Nash Equilibria (FNE)*** if
1) all routes with human traffic have latency $\ell_{m_{\mathrm{EQ}}} \leq \ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i) \; \forall i \in [N]$ and
2) for any $\ell \geq 0$, a volume of at least $\varphi(\ell/\ell_{m_{\mathrm{EQ}}})\lambda^{\mathrm{a}}$ autonomous traffic experiences a latency less than or equal to ℓ. Note that it is sufficient to check this condition for $\ell = \ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i)$ for all i.

We denote the set of routings at Flexible Nash Equilibria with demand $(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}})$, equilibrium latency $\ell_{m_{\mathrm{EQ}}}$, and flexibility profile φ as FNE$(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}}, \ell_{m_{\mathrm{EQ}}}, \varphi)$. The set of ***Best-case Flexible Nash Equilibria (BFNE)*** is the subset of FNE with

routings that minimize total latency. Note that as in Theorem 1, we use m_{EQ}^* to denote the road with longest free-flow latency that contains selfish vehicle flow in the best routing within the considered set of equilibria. We defer the proof of the following to the appendix.

Theorem 2. *For any given routing in the set of BFNE, there exist a longest equilibrium road m_{EQ}^* and a longest used road m_{ALL}^* with $m_{\mathrm{EQ}}^* \leq m_{\mathrm{ALL}}^*$, such that:*
1) *roads with index less than m_{EQ}^* are congested,*
2) *roads with index greater than m_{EQ}^* are in free-flow,*
3) *roads with index greater than m_{EQ}^* and less than m_{ALL}^* have maximum flow.*

Remark 1. *Note that, unlike in BNE, road m_{EQ}^* will not necessarily be in free-flow and m_{EQ}^* is not necessarily the minimum index such that all selfish traffic can be feasibly routed at Nash Equilibrium [30]. Further, different elements of the set BFNE can have different indices for longest equilibrium and longest used road.*

Finding the Best-case Nash Equilibria. In general, the Nash Equilibrium constraint is a difficult combinatorial constraint. Theorem 1 however states that we can characterize the congestion profile of the roads by finding the minimum free-flow road such that Nash Equilibrium can be feasibly achieved. This is formalized as follows: find the minimum m_{EQ} such that NE$(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}}, m_{\mathrm{EQ}})$ is nonempty:

$$m_{\mathrm{EQ}}^* = \underset{m_{\mathrm{EQ}} \in [N]}{\arg\min} \; a_{m_{\mathrm{EQ}}} \text{ s.t. } \mathrm{NE}(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}}, m_{\mathrm{EQ}}) \neq \emptyset \text{ , then}$$
$$\mathrm{BNE}(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}}) \; \subseteq \; \mathrm{NE}(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}}, m_{\mathrm{EQ}}^*). \tag{4}$$

Theorem 3. (4) *can be solved in $O(N^4)$ time.*

We defer the proof to the appendix.
Finding the Best-case Flexible Nash Equilibria. To find an element of the BFNE, we need to solve:

$$\underset{\substack{m_{\mathrm{EQ}} \in [N], \; \hat{\ell}_0 \in [a_{m_{\mathrm{EQ}}}, a_{m_{\mathrm{EQ}}+1}), \\ (\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) \in \mathrm{FNE}(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}}, \hat{\ell}_0, \varphi)}}{\arg\min} \quad C(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) . \tag{5}$$

As demonstrated in [30], the longest equilibrium road is no longer the road with lowest free-flow latency such that the routing is feasible, as was the case in BNE. Further, road m_{EQ} may not be in free-flow in the set of BFNE. However, we do know that for a fixed m_{EQ}, the latency on road m_{EQ} which minimizes cost, subject to feasibility constraints, is one of a finite number of options.

Theorem 4. *Finding a solution to* (5) *is equivalent to finding a routing in the set of BFNE, if any exist. Further,* (5) *can be solved in $O(|K|N^5)$ time, where $|K|$ is the number of flexibility levels of autonomous vehicle users.*

Proof. First, Definition 2 implies that the latency on the longest equilibrium road m_{EQ}, which we denote $\hat{\ell}_0$, must be less than that of road $m_{\mathrm{EQ}}^* + 1$. This, with the definition of BFNE imply that (5) solves for an element of the BFNE. Now, note that for a given $\hat{\ell}_0$, the optimal routing will maximize the autonomous flow on roads $[m_{\mathrm{EQ}}]$. We show that this can be computed in $O(N^3)$ time, and the optimal allocation of the

remaining autonomous flow can be computed in $O(N)$ time. Next, we note that subject to feasibility, the social cost of a routing decreases monotonically with $\hat{\ell}_0$. In light of this, we show that there are a maximum of $k|N|$ critical points to check for feasibility when searching for the optimal $\hat{\ell}_0$.

We now show that given $\hat{\ell}_0$, the latency on road m_{EQ}, computing the optimal flow on roads $[m_{\mathrm{EQ}}]$ can be done in $O(N^3)$ time. For a given $\hat{\ell}_0$, the optimal routing will fit as much autonomous flow as possible on roads $[m_{\mathrm{EQ}}]$. Let $\boldsymbol{f}^{\mathrm{h,EQ}}$ and $\boldsymbol{f}^{\mathrm{a,EQ}}$ denote the elements of the regular and autonomous routings $\boldsymbol{f}^{\mathrm{h}}$ and $\boldsymbol{f}^{\mathrm{a}}$ that correspond to flows on the roads $[m_{\mathrm{EQ}}]$, as with $\boldsymbol{s}^{\mathrm{EQ}}$. Then, solve

$$
\underset{\boldsymbol{f}^{\mathrm{h,EQ}},\boldsymbol{f}^{\mathrm{a,EQ}}\in\mathbb{R}_{\geq0}^{|[m_{\mathrm{EQ}}]|},s_{m_{\mathrm{EQ}}}\in\{0,1\}}{\arg\max}\sum_{i\in[m_{\mathrm{EQ}}]}f_i^{\mathrm{a}} \tag{6}
$$

$$
\text{s.t.}\sum_{i\in[m_{\mathrm{EQ}}]}f_i^{\mathrm{h}}=\lambda^{\mathrm{h}},\ \sum_{i\in[m_{\mathrm{EQ}}]}f_i^{\mathrm{a}}\leq\lambda^{\mathrm{a}},
$$

$$
\forall i\in[m_{\mathrm{EQ}}]\ f_i^{\mathrm{h}}\geq0,f_i^{\mathrm{a}}\geq0,
$$

$$
\forall i\in[m_{\mathrm{EQ}}-1]\ \ell_i(f_i^{\mathrm{h}},f_i^{\mathrm{a}},1)=\hat{\ell}_0,
$$

$$
\ell_{m_{\mathrm{EQ}}}(f_{m_{\mathrm{EQ}}}^{\mathrm{h}},f_{m_{\mathrm{EQ}}}^{\mathrm{a}},s_{m_{\mathrm{EQ}}})=\hat{\ell}_0.
$$

Using similar reasoning as in Theorem 3, this can be formulated as a linear program and therefore can be solved with a computational complexity $O(N^3)$ [31].

Having computed the optimal routing on roads $[m_{\mathrm{EQ}}]$, we now consider an optimization which computes the resulting optimal index of the longest used road in order to fit the autonomous flow, and as a result, the optimal routing of the remaining autonomous flow. This follows from Theorem 2:

$$
\underset{j\in[N]\setminus[m_{\mathrm{EQ}}-1]}{\arg\min}\ j
$$

$$
\text{s.t.}\quad\sum_{i\in[m_{\mathrm{EQ}}]}f_i^{\mathrm{a}}+\sum_{i\in[j]\setminus[m_{\mathrm{EQ}}]}\bar{F}_i(1)\geq\lambda^{\mathrm{a}}, \tag{7}
$$

which requires computations of order $O(N)$.

We temporarily restrict our attention to the case in which autonomous users have a uniform autonomy level. We wish to optimize over the following decision variables:

$m_{\mathrm{EQ}}\in[N]$ longest equilibrium road?

$m_{\mathrm{ALL}}\in[m]\setminus[m_{\mathrm{EQ}}-1]$ longest used road?

$\hat{\ell}_0\in[a_{m_{\mathrm{EQ}}},a_{m_{\mathrm{EQ}}+1})$ equilibrium latency?

$\boldsymbol{f}^{\mathrm{h}},\boldsymbol{f}^{\mathrm{a}}\in\mathbb{R}_{\geq0}^{n},\boldsymbol{s}\in\{0,1\}^{n}$ actual routing?

The objective function to be minimized is aggregate latency, which, using the Theorem 2, can be formulated as follows:

$$
\hat{\ell}_0\sum_{i\in[m_{\mathrm{EQ}}]}(f_i^{\mathrm{h}}+f_i^{\mathrm{a}})+\sum_{i\in[m_{\mathrm{ALL}}-1]\setminus[m_{\mathrm{EQ}}]}a_i\bar{F}_i(1)+a_{m_{\mathrm{ALL}}}\Big(\lambda^{\mathrm{a}}-\sum_{i\in[m_{\mathrm{ALL}}-1]}f_i^{\mathrm{a}}\Big)
$$

$$
\text{s.t. } (\boldsymbol{f}^{\mathrm{h,EQ}},\boldsymbol{f}^{\mathrm{a,EQ}},\boldsymbol{s}^{\mathrm{EQ}})\in(6) \tag{8}
$$

$$
m_{\mathrm{ALL}}=(7) \tag{9}
$$

$$
\frac{1}{\lambda^{\mathrm{a}}}\Big(\sum_{i\in[m_{\mathrm{EQ}}]}f_i^{\mathrm{a}}+\sum_{i\in[j]\setminus[m_{\mathrm{EQ}}]}\bar{F}_i(1)\Big)\geq\varphi\Big(\frac{a_i}{\hat{\ell}_0}\Big)\ \forall j\in[N]\setminus[m_{\mathrm{EQ}}-1] \tag{10}
$$

For a given $\hat{\ell}_0$, the optimal routing maximizes the autonomous flow on roads $[m_{\mathrm{EQ}}]$, yielding (8). The remaining

autonomous flow is routed as in (9). Finally, (10) ensures that no one is more flexible than they wish.

To solve this, recall that in the case of uniform flexibility,

$$
\varphi(\kappa)=\begin{cases}0 & 0\leq\kappa\leq\kappa_0\\1 & \kappa>\kappa_0\ .\end{cases}
$$

Further, the volume of autonomous flow that can fit on roads $[m]$ increases with decreasing ℓ_m. Because of this, we can restrict our search of $\hat{\ell}_0$ to critical points of the function φ:

$$
\hat{\ell}_0\in\{a_{m_{\mathrm{EQ}}}\}\cup\{\frac{a_i}{\kappa_0}:i\in[m_{\mathrm{ALL}}]\setminus[m_{\mathrm{EQ}}],a_{m_{\mathrm{EQ}}}<\frac{a_i}{\kappa_0}<a_{m_{\mathrm{EQ}}+1}\},
$$

which is a set with maximum cardinality N. Therefore, we can find the BFNE via the following algorithm:

1) Enumerate through all possible values of m_{EQ} (N possibilities).
2) For each possible m_{EQ}, enumerate through all possible values of $\hat{\ell}_0$ (N possibilities).
3) For each combination of m_{EQ} and $\hat{\ell}_0$, find the optimal routing on roads $[m_{\mathrm{EQ}}]$ via (6) (order N^3), and find m_{ALL} and the optimal routing of autonomous vehicles on the remaining roads via (7) (order N). As these are sequential, this step requires computations of order $O(N^3)$.

All together, this requires computations of order $O(N^5)$.

Now consider that autonomous vehicles have nonuniform flexibility levels. Then,

$$
\varphi(\kappa)=\begin{cases}0 & 0\leq\kappa\leq\kappa_0\\\varphi_0 & \kappa_0<\kappa\leq\kappa_1\\\varphi_1 & \kappa_1<\kappa\leq\kappa_2\\\dots\\1 & \kappa>\kappa_{|K|-1}\ .\end{cases}
$$

We therefore must search over $\hat{\ell}_0$ in the following set:

$$
\hat{\ell}_0\in\{a_{m_{\mathrm{EQ}}}\}\cup
$$

$$
\{\frac{a_i}{\kappa_j}:i\in[m_{\mathrm{ALL}}]\setminus[m_{\mathrm{EQ}}],j\in K,a_{m_{\mathrm{EQ}}}<\frac{a_i}{\kappa_j}<a_{m_{\mathrm{EQ}}+1}\},
$$

which has maximum cardinality $|K|N$, bringing the total computation complexity to $O(|K|N^5)$. $\qquad\square$

IV. INCENTIVIZING FLEXIBILITY

We next present a pricing mechanism to attain the same benefits of flexible behavior through financial incentives in the case of elastic demands. We first explain the human choice model that formalizes how autonomous service users choose between a variety of price and latency pairs depending on their value of time, encoded by their *reward functions*. We then formulate a mathematical optimization problem that aims to find an optimal trade-off between high road usage and low average travel time. We then explain the data-efficient algorithm that we propose for learning human reward functions.

A. Autonomous Service User Choice Model

How users choose between a variety of price and latency pairs depends on their valuation of time and money. Without knowing this choice model we cannot plan vehicle flows and will not be able to ensure the resulting configuration matches

our vehicle flow models for the roads. Also, since different populations may have different valuations, we need to learn this tradeoff for our population so we can estimate how many people will choose which option.

To untangle these constraints, we describe how autonomous service users make routing decisions (the model for human drivers is provided in Sec. II-C). Though human drivers are motivated directly only by latency, autonomous service users experience cost in both latency and the price of the ride. We model the users as having some underlying reward function, which is parameterized by their time/money tradeoff, as well as the desirability of the option of traveling by some other means such as walking or public transit. We assume that a strictly dominated option in terms of latency and price is completely undesirable. We formally define this set

$$D = \{i \in [N] \mid (p_i > p_{i'} \wedge \ell_i \geq \ell_{i'}) \vee (p_i \geq p_{i'} \wedge \ell_i > \ell_{i'})$$
$$\text{for some } i' \in [N]\} \,,$$

where p_i is the monetary cost of route i. We also define the set of undominated roads, $\bar{D} = [N] \backslash D$. We model the reward function of user j for choosing road i as follows, where we assume choosing road 0 denotes declining the service:

$$r_j(\boldsymbol{\ell}, \mathbf{p}, i) =$$
$$\begin{cases} -\omega_{j_1}\ell_i - \omega_{j_2}p_i & \text{if } i \in \bar{D}, \\ -\infty & \text{if } i \in D, \\ -\zeta_j\ell^w & \text{if } i = 0 \text{ (user } j \text{ declines the service)} \,, \end{cases}$$

with the following nonnegative parameters: $\boldsymbol{\ell}$ denotes the vector of road latencies, $\boldsymbol{p}$ denotes prices, $\boldsymbol{\omega}_j = \begin{bmatrix} \omega_{j_1} & \omega_{j_2} \end{bmatrix}^\top$ characterizes the users' time/money tradeoff and ζ_j specifies their willingness to use an alternative option with latency ℓ^w, which could be walking, biking, or public transportation.

We do not assume users are simple reward maximizers. Rather, we adopt the multinomial logit model [4] for the probability with which users choose each option:

$$P(\text{user } j \text{ chooses route } i) \propto \exp\left(r_j(\boldsymbol{\ell}, \boldsymbol{p}, i)\right) \quad (11)$$

for all $i \in [N] \cup \{0\}$.

In order to determine the optimal pricing, we want to know how many users will choose each route as a function of the route prices and latencies. To this end, we define $q_i(\boldsymbol{\ell}, \boldsymbol{p})$ as the expected fraction of autonomous service users that will choose route i. If the parameter distribution for autonomous service users is $g(\boldsymbol{\omega}, \zeta)$, then

$$q_i(\boldsymbol{\ell}, \boldsymbol{p}) = \int_0^\infty \int_0^\infty \int_0^\infty g(\boldsymbol{\omega}, \zeta)P(i \mid \boldsymbol{\omega}, \zeta)d\omega_1 d\omega_2 d\zeta \,.$$

where $P(i \mid \boldsymbol{\omega}, \zeta)$ is the probability that a user with reward function parameters $(\boldsymbol{\omega}, \zeta)$ will choose road i. This expression relate prices and latency to human choices, enabling us to determine the prices that will maximize the social objective. This will be important in constraining the optimization to only consider latency/price options that correspond to the desired vehicle flows.

B. Solution Method

Problem Formulation. We now formulate the problem where we have an indirect control of the autonomous cars' routing through pricing, and the demand of autonomous service is elastic. Because of this elasticity, we cannot just minimize the average latency, which would result in extremely high prices to keep autonomous service users off the network. Hence, we consider an objective that is a combination of maximizing road usage and minimizing average travel time.[1] We parameterize this tradeoff with parameter $\theta \geq 0$ in the cost function

$$J(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) = \frac{\sum_{i \in [N]} (f_i^{\mathrm{h}} + f_i^{\mathrm{a}})\ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i)}{\sum_{i \in [N]} (f_i^{\mathrm{h}} + f_i^{\mathrm{a}})} - \theta \sum_{i \in [N]} (f_i^{\mathrm{h}} + f_i^{\mathrm{a}}) \,.$$

Our control variables in this optimization are the *latency* on each road and the *price* offered to the users for traveling on each route. However, we cannot arbitrarily choose prices and latencies – we need to respect

1) *the characteristics of the roads*, in terms of how the flow demand for a road corresponds to the latency on that road (Section II-A), and

2) *how people make decisions*, making sure that the number of people who choose each option corresponds to the latency of the roads described in the options (Section IV-A).

Moreover, we want to be fair; so we must offer the same pricing and routing options to all autonomous service users.

Given $\mathbf{q}(\boldsymbol{\ell}, \boldsymbol{p})$, the cost function $J(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s})$, inelastic flow demand of human drivers λ^{h}, and elastic demand of autonomous users λ^{a}, we are ready to formulate the planning optimization. The most straightforward way is to optimize jointly over $\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{\ell}$ and $\boldsymbol{p}$. However, $\boldsymbol{\ell}$ is fully defined by $\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}$ and $\boldsymbol{s}$. Hence, instead of $\boldsymbol{\ell}$, we use $\boldsymbol{s}$, which will help us solve this nonconvex optimization. The problem is then formulated as:

$$\min_{\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{p} \in \mathbb{R}_{\geq 0}^N, k \in [N], s_{k:N} \in \{0,1\}^{N-k+1}} J(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) \quad (12)$$

$$\text{s.t. } \sum_{i \in [N]} f_i^{\mathrm{h}} = \lambda^{\mathrm{h}} \quad (13)$$

$$f_i^{\mathrm{a}} = \lambda^{\mathrm{a}} q_i(\boldsymbol{\ell}(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}), \boldsymbol{p}), \forall i \in ([N] \backslash [k]) \cup D \quad (14)$$

$$\sum_{i \in [k] \backslash D} f_i^{\mathrm{a}} = \lambda^{\mathrm{a}} \sum_{i \in [k] \backslash D} q_i(\boldsymbol{\ell}(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}), \boldsymbol{p}) \quad (15)$$

$$a_k \leq \ell_k(f_k^{\mathrm{h}}, f_k^{\mathrm{a}}, s_k) \leq a_{k+1} \quad (16)$$

$$f_i^{\mathrm{h}} = 0, \forall i \in [N] \backslash [k] \quad (17)$$

$$\ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, 1) = \ell_k(f_k^{\mathrm{h}}, f_k^{\mathrm{a}}, s_k), \forall i \in [k-1] \quad (18)$$

$$f_i^{\mathrm{h}} + f_i^{\mathrm{a}} \leq \bar{F}\left(f_i^{\mathrm{a}}/(f_i^{\mathrm{a}} + f_i^{\mathrm{h}})\right), \forall i \in [N] \quad (19)$$

$$\sum_{i \in [N]} (f_i^{\mathrm{a}} p_i - f_i^{\mathrm{a}} d_i c) \geq \bar{P} \quad (20)$$

with $s_{1:k-1} = 1$ due to selfishness of human-driven vehicles. Here, k is the longest road with human-driven vehicle flow, $\bar{P}$ is the minimum profit per unit time for the autonomous service provider and c is the constant fuel cost per unit length. We can describe the constraints as follows.

[1]Some other works, e.g. [32], use the *social welfare* objective. While the two objectives have many similarities, we cannot directly adopt it here as we have heterogeneous users who have different price-latency valuations.

13) The human-driven vehicle flow demand is fulfilled.
14) Autonomous flow will be distributed into the roads in $([N] \setminus [k]) \cup D$ based on the choice model described in the preceding section.
15) Total autonomous flow in $[k] \setminus D$ will satisfy the user choices, but can be distributed arbitrarily as the roads have the same latency and price.
16) The "longest equilibrium road" has latency on the given interval of free flow latency.
17) Human-driven cars are selfish, i.e. no human-driven car will experience higher latency than the road k.
18) The congested roads have the same latency as the "longest equilibrium road".
19) The maximum capacities of the roads are respected.
20) The minimum profit per unit time is satisfied.

We can further improve the search space by relying on the heuristic that the roads that are not used by the human-driven vehicles will be in free-flow, i.e. $s_{k+1:N} = 0$. While we do not have a proof for the conditions that lead to this, we also note constructing counterexamples seems to require extremely careful design, which suggests the heuristic holds in general. Furthermore, the following theorem shows we could also set $s_k = 0$ under an additional assumption.

Theorem 5. *Assume $\omega_2 > 0$ for all users. Then there exists a free-flow road k in the optimal solution to the problem such that $\ell_i = \ell_k$ for $\forall i \in [k]$, and $f_i^h = 0$ for $\forall i \in [N] \setminus [k]$ as long as the optimization is feasible.*

We defer the proof to the appendix.

Generalizations. We assumed all autonomous cars are controlled by a centralized social planner. To extend our framework to scenarios where this is not the case and the social planner has the control over a fraction of autonomous cars, we can simply do the following modifications: The optimization will also be over $\boldsymbol{f}^{\mathrm{b}} \in \mathbb{R}_{\geq 0}^N$, which will now represent the autonomous flow that does *not* belong to the planner. Add the corresponding constraint of (13) for f_i^{b}. Similar to (17), $f_i^{\mathrm{b}} = 0$ for $i \in [N] \setminus [k]$ due to selfishness. Also, replace f^{a}'s in (16), (18) and (19) with $f^{\mathrm{a}} + f^{\mathrm{b}}$ with appropriate subscripts. These simple modifications enable a more general use.

Solving the Optimization. After learning the distribution $g(\boldsymbol{\omega}, \zeta)$ (described below), we first take M samples $(\bar{\boldsymbol{\omega}}, \bar{\zeta}) \sim g(\boldsymbol{\omega}, \zeta)$. Using these samples, we approximate the expected fraction of autonomous users that will choose route i as:

$$q_i(\boldsymbol{\ell}, \boldsymbol{p}) \doteq \frac{1}{M} \sum_{\bar{\boldsymbol{\omega}}, \bar{\zeta}} P(i \mid \bar{\boldsymbol{\omega}}, \bar{\zeta})$$

where $\doteq$ denotes asymptotic equality as $M \to \infty$.

We then locally solve the nonconvex planning optimization using interior point algorithms [33] with 100 random initial points for each run to get closer to global optimum.

Data-Efficient Learning of Human Reward Functions. While the routes in a specific network can be fully modeled with the physical properties and the speed limits, user's decision models (parameterized by $(\boldsymbol{\omega}, \zeta)$) must be learned in a data-driven way. The parameters might be different among the users. While a business executive might prefer paying extra

to reach their company quickly, a graduate student may decide to go to the lab a little later in order to save a few dollars. Therefore, we have to learn personalized parameters $\boldsymbol{\omega}$ and ζ.

We learn the parameters from users' previous choices, which is known as *preference-based learning*. If user j chooses from a variety of options, the user's choice gives us a noisy estimate of which road $i \in [N] \cup \{0\}$ maximizes their reward function $r_j(\boldsymbol{\ell}, \boldsymbol{p}, i)$. We could start from either uniform priors or priors informed by domain knowledge, then sample from the distribution $g(\boldsymbol{\omega}, \zeta)$.

However, a major drawback of doing so is how quickly we learn the user preferences. Preference-based learning suffers from the small amount of information that each query carries. For example, if we show 4 options to a user (including the option to decline the service), then the maximum information we can get from that query is only 2 bits. To tackle this issue, previous works pose the query synthesis problem as an optimization and maximize a measure of the expected change in the learned distribution after the queries [6].

While those works focus on pairwise queries, in this case we expect to pose several route options to the users and therefore need more general query types. By using these general queries which offer a variety of routes with varied latency and price, we can consider various ways of using this learning framework to learn the human preferences.

- We could do a user study on a few people to learn a good prior distribution.
- We could use an exploration/exploitation strategy if we are allowed to break the fairness constraint a few times for some small portion of the users. This could be made through user-specific promotions; for each user we may either choose to use the learned model or to offer special rates that would help us profile the user better.
- We could do an initial profiling study for each new user.

To implement any of these options, we formulate the following active learning optimization. First, we discuss the general preference-based learning framework. Given the data from previous choices of user j, which we denote as $\mathcal{D}_j$, we formalize the probability of $(\boldsymbol{\omega}_j, \zeta_j)$ being true parameters for that user as follows:

$$P(\boldsymbol{\omega}_j, \zeta_j \mid \mathcal{D}_j) \propto P(\boldsymbol{\omega}_j, \zeta_j) \prod_m P(\mathcal{D}_{j_m} \mid \boldsymbol{\omega}_j, \zeta_j)$$

where $\mathcal{D}_{j_m}$ denotes the road user j chose in their m^{th} choice (with $\mathcal{D}_{j_m} = 0$ meaning that the user declined the service and preferred the alternative option). The relation is due to the assumption that the users' choices are conditionally independent from each other given the reward function parameters.

The second term comes from the human choice model. For the prior, we can use a uniform distribution over nonnegative parameters. The prior may be crucial especially when we do not have enough data for a new user. In such settings, we incorporate domain knowledge to start with a better prior.

We then use this unnormalized $P(\boldsymbol{\omega}_j, \zeta_j \mid \mathcal{D}_j)$ to obtain the samples of $(\boldsymbol{\omega}_j, \zeta_j)$ using Metropolis-Hastings algorithm. Doing this for each user, which can be easily parallelized, we directly obtain the samples $(\bar{\boldsymbol{\omega}}, \bar{\zeta}) \sim g(\boldsymbol{\omega}, \zeta)$.

Next we formulate the active learning framework, which is needed so that it will not take an excessive number of queries to learn human preferences. For this, we want to maximize the expectation of the difference between the prior and the unnormalized posterior:

$$\text{query}_m^* = \underset{\text{query}_m}{\arg\max}\, \mathbb{E}_{\mathcal{D}_{j\,m}} \Big[P(\boldsymbol{\omega}_j, \zeta_j | \mathcal{D}_{j\,1:m-1}) -$$
$$P(\boldsymbol{\omega}_j, \zeta_j | \mathcal{D}_{j\,1:m-1}) P(\mathcal{D}_{j\,m} | \boldsymbol{\omega}_j, \zeta_j, \mathcal{D}_{j\,1:m-1}) \Big]$$
$$= \underset{\text{query}_m}{\arg\min}\, \mathbb{E}_{\mathcal{D}_{j\,m}} \Big[P(\mathcal{D}_{j\,m} | \boldsymbol{\omega}_j, \zeta_j, \mathcal{D}_{j\,1:m-1}) \Big]$$

As we will use the sampled $(\boldsymbol{\omega}_j, \zeta_j)$ to compute the probabilities of road choices, we can write the optimization as:

$$\text{query}_m^* \doteq \underset{\text{query}_m}{\arg\min}\, \mathbb{E}_{\mathcal{D}_{j\,m}} \left[\sum_{\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j} P(\mathcal{D}_{j\,m} | \bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j, \mathcal{D}_{j\,1:m-1}) \right]$$

where we have M samples denoted as $(\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j)$, and the term $1/M$ is canceled. Using the law of total probability,

$$P(\mathcal{D}_{j\,m} | \mathcal{D}_{j\,1:m-1}) \doteq \frac{1}{M} \sum_{\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j} P(\mathcal{D}_{j\,m} | \bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j, \mathcal{D}_{j\,1:m-1})$$

which leads to the following optimization for finding query_m^*:

$$\underset{\text{query}_m}{\arg\min} \sum_{\mathcal{D}_{j\,m}} P(\mathcal{D}_{j\,m} | \mathcal{D}_{j\,1:m-1}) \sum_{\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j} P(\mathcal{D}_{j\,m} | \bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j, \mathcal{D}_{j\,1:m-1})$$

$$\doteq \underset{\text{query}_m}{\arg\min} \sum_{\mathcal{D}_{j\,m}} \left(\sum_{\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j} P(\mathcal{D}_{j\,m} | \bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j) \right)^2$$

We can easily compute this objective value for any given query_m. This optimization is nonconvex due to the human choice model. As in previous works, we assume local optima is good enough [6]. We then use a Quasi-Newton method (L-BFGS [34]) to find the local optima, and we repeat this for 1000 times starting from random initial points.

V. USER EXPERIMENTS

To validate our framework, we conducted different simulations and a user study approved by Stanford University's Research Compliance Office.

Hypotheses. We test three hypotheses that together suggest our framework successfully reduces traffic congestion through pricing, after it learns the humans' choice models:

H1: Our active learning algorithm can learn the autonomous service user preferences in a data-efficient way.

H2: Our planning optimization reduces the overall latency by creating flexible behavior through pricing.

H3: When used by humans, the overall framework works well and is advantageous over inflexible algorithms.

Implementation Details. In the planning optimization, we used the heuristic $s_{k+1:N} = 0$. We assumed the only alternative for autonomous service users is walking. We set $c = 6 \times 10^{-5}$ USD/meter, $\mathbf{b} = 1$. We assumed the regular and autonomous cars keep a 2-second and 1-second headway distance with the leading car, respectively. The length of the cars is 5 meters, and the minimum gap between two cars is 2 meters.

Experiments and Analyses. To validate **H1**, we simulated 5 autonomous service users with different preferences. We tested our active learning framework by asking two sets of 200 queries, each of which consisted of 4 road options, similar to Fig. 6, and a walking option. The queries were generated actively in the first set and randomly in the second. After each query, we recorded the sample $(\bar{\boldsymbol{\omega}}, \bar{\zeta})$ which has the highest likelihood as our estimates.

Fig. 4 shows how the estimates evolved within active learning setting for one of the users. All values are overestimated initially. Intuitively, this is because getting noiseless responses has higher likelihood in the beginning. As we query more, accepting some of the responses as noisy maximizes the likelihood. Therefore, the values start decreasing.
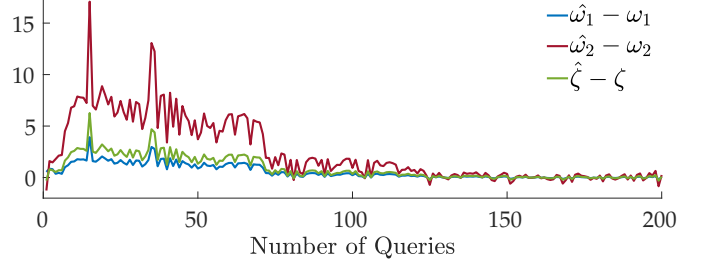

Fig. 4: The errors of the reward function estimates are shown with varying number of queries. $\hat{\omega}_1$, $\hat{\omega}_2$ and $\hat{\zeta}$ represent the estimates.

Another important observation is that the estimates of the parameters increase and decrease together even in the early iterations. This suggests we are able to learn the ratio between the parameters, e.g. ω_1/ω_2 very quickly. To check this claim, we used the following error metric:

$$e_{x,y} = \|x/y - \hat{x}/\hat{y}\|_1$$

where $x, y \in \{\omega_1, \omega_2, \zeta\}$ and $\hat{x}, \hat{y}$ represent the corresponding estimates. Fig. 5 shows how this error decreases with increasing number of queries. It also shows how active querying enables data-efficient learning compared to the random querying baseline. We are able to learn the relationship between parameters even under 20 queries with active learning. All these results strongly support **H1**.
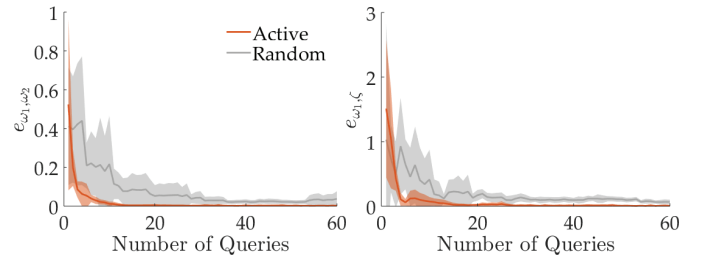

Fig. 5: The error metric is averaged over 5 different reward functions.

The fact that we are able to learn the ratios implies we can estimate which road the user is most likely to choose. We will only be unsure about how noisy the user is if the parameter estimates did not converge yet. Therefore, we can still use the estimates for our planning optimization even when we have small number of queries.

To validate **H2**, we use the road network from [30] and the equilibria benchmarks we developed: NE, BNE and BFNE with full flexibility. Here, we give the average latencies for the 4-road network from that study which we visualize in Fig. 6,
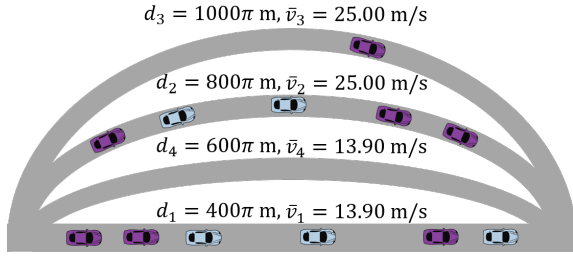
Fig. 6: The 4-road network from [30]. The roads are not to the scale and ordered with respect to the free-flow latencies.

and where $\lambda^h = 0.4, \lambda^a = 1.2$ cars per second: (NE: 400.00 sec, BNE: 125.66 sec, BFNE (full flexibility): 102.85 sec).

We then assumed we perfectly learned the preferences of the 5 simulated users. We ran the planning optimization with $\bar{P} = 0$ and 3 different θ to show the trade-off. The results are summarized in Table I.

TABLE I: Results of Routing Simulation

θ	Avg. Latency (seconds)	Flow (cars/second)
1	90.41	0.4412
20	97.03	1.2746
10^6	111.28	1.5964

It can be seen we can adjust the trade-off between average latency and the served flow by tuning θ. Also, given the human preferences, even when we served (almost) all of autonomous demand, our framework outperforms BNE. This shows its effectiveness on creating flexibility and supports **H2**.

For **H3**, we recruited 21 subjects (9 female, 12 male) with an age range from 19 to 60. In the first phase of the experiment, each participant was asked 40 actively synthesized queries (4 roads + 1 walking option). We then used their responses to get the maximum likelihood sample $(\bar{\omega}, \bar{\zeta})$. Afterwards, we designed 5 different road networks each with 4 different roads and an additional route where people may choose to walk. The 5 different networks cover a range of different road lengths from 1.8 kilometers to 78 kilometers. For each network, we also set different $\theta, \lambda^h, \lambda^a$, and $\bar{P}$. By assuming all autonomous flow is in the service of these 21 subjects, or of the groups that match with their preferences, we locally solved the planning optimization to get the pricing scheme for each traffic network. We refer to these results as *anticipated* values.

In the second phase of the user study, we presented the route-price pairs and the walking option to the same 21 subjects. For each of the 5 networks, they picked the option they would prefer. Using these responses, we allocated the autonomous flow into the roads. However, it is technically possible that more users select a road than its maximum flow. To handle such cases, we assumed extra flow is transferred to the roads with smaller latencies without making the users pay more. If that is not feasible, extra flow is transferred to the slower roads, without any discount. While these break the fairness constraint, it rarely happens and affects only a very small portion of the flow. After autonomous flows are allocated, human driven cars selfishly chose their routes in a way to minimize the overall average latency. We refer to the results of this allocation as *actual* values.

Table II compares the anticipated and the actual values. We report latencies in seconds, flows in cars per second, and profit is a rate value with the unit USD per second. In order to show

how our framework incentivizes flexible behavior, we also added other benchmarks: two where the same flow as actual flow is routed (BNE1 and BFNE1) and two where all flow demand is routed, i.e. no walking option (BNE2 and BFNE2). While BNE and BFNE assume completely inelastic demand, using them as benchmarks under both the actual flow and the all flow demand gives us insights about the success of our framework, because the allocation of vehicles under the actual flow is comparable to the elastic demand case.

It can be seen that there is generally an alignment between the anticipated and actual values. While the mismatch may be further reduced by doing more queries or having more users, the difference with BNE1 is significant. In all cases, our framework achieved to incentivize flexibility, which yielded lower average latencies compared to BNE1. Especially in Case 3 and Case 5, our framework approximately halved the average latency compared to the best Nash equilibria. In fact, our framework achieved a latency that is close to the lower bound set by BFNE1. We visualize these in Fig. 7. Furthermore, our framework successfully reduced flow demand when satisfying the full demand under selfishness is impossible (Case 1).
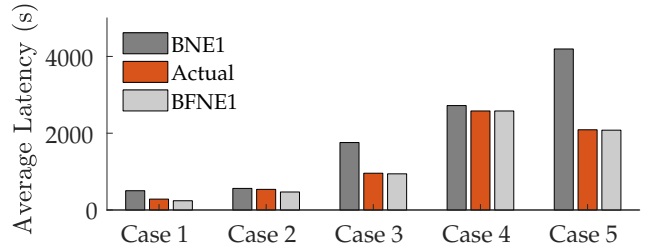

Fig. 7: The comparison of the actual results and BNE1, both of which allocate the same amount of flow.

One caveat is the small amount of actual flow in Case 1, which also caused an important profit loss. This is because the roads are relatively shorter, and most users preferred walking over paying for an autonomous car. Our framework could not predict this, because the learned reward functions failed to accurately estimate the probabilities.

VI. CONCLUSION

In this work we address the efficiency of traffic networks with mixed autonomy. We develop a method of pricing rides with autonomous vehicles such that when a population chooses from these route and price options, and the human drivers choose the quickest routes available to them, the objective of decreasing travel latency and increasing road usage is achieved. To do so, we model how people choose between different route options with varying prices and latencies. Moreover, we develop a method for actively learning the parameters that describe the preferences of a population of users. We develop theoretical results which we use to gauge the performance of our algorithm, and conduct a user study showing that our method of parameterizing and actively learning the preferences of a human population is effective.

A wide horizon for further research remains. One could relax the assumption that the reward functions are linear to improve the prediction accuracy [35] and optimize for information gain in the active learning scheme, which can yield better data efficiency [36]. Another direction is to consider

TABLE II: Results of Real-User Experiments

	Anticipated			Actual Pricing Scheme			BNE1	BFNE1	Full	BNE2	BFNE2
	Flow	Av. Latency	Profit	Flow	Av. Latency	Profit	Av. Latency	Av. Latency	Flow	Av. Latency	Av. Latency
Case 1	1.6622	338.10	11.11	1.0429	283.30	3.85	501.97	240.80	1.9000	Infeasible	421.45
Case 2	1.5592	510.97	4.50	1.5048	537.30	3.95	562.20	468.79	1.6000	562.20	474.35
Case 3	1.2677	921.06	4.59	1.3000	957.71	3.99	1756.89	941.81	1.3000	1756.89	941.81
Case 4	1.5538	2576.36	40.01	1.6000	2579.52	49.33	2720.00	2579.33	1.6000	2720.00	2579.33
Case 5	1.2383	2083.42	48.48	1.2048	2089.53	46.32	4194.26	2079.15	1.3000	4194.26	2183.56

For a fixed flow, BFNE serves as a lower bound on the latency. Achieving lower latency than BNE means we successfully incentivize flexible behavior.

an elastic demand for human drivers, as well as a variety of nondriving options including walking, biking, or taking the bus. In that case, ζ will have a multimodal distribution; we then need to learn the mixture.

More broadly, one can look at more general network topologies – if route pricing is computed only with local information on each edge, pricing can make congestion worse when users have different price sensitivities [37]. One can also expand the model to include the role of information in decision making, as well as biases such as risk aversion. Through these future directions we can ensure the efficient operation of transportation networks that include autonomous vehicles.

REFERENCES

[1] D. Schrank, B. Eisele, T. Lomax, and J. Bak, "Urban mobility scorecard," 2015.

[2] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," in *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, 1999.

[3] N. Mehr and R. Horowitz, "Can the presence of autonomous vehicles worsen the equilibrium state of traffic networks?" in *IEEE Conference on Decision and Control (CDC)*, 2018.

[4] M. E. Ben-Akiva, S. R. Lerman, and S. R. Lerman, *Discrete choice analysis: theory and application to travel demand*. MIT press, 1985, vol. 9.

[5] S. Dafermos, "Traffic equilibrium and variational inequalities," *Transportation science*, 1980.

[6] D. Sadigh, A. D. Dragan, S. Sastry, and S. A. Seshia, "Active preference-based learning of reward functions," in *Robotics: Science and Systems (RSS)*, 2017.

[7] J. Lioris, R. Pedarsani, F. Y. Tascikaraoglu, and P. Varaiya, "Platoons of connected vehicles can double throughput in urban roads," *Transportation Research Part C: Emerging Technologies*, 2017.

[8] A. Askari, D. A. Farias, A. A. Kurzhanskiy, and P. Varaiya, "Effect of adaptive and cooperative adaptive cruise control on throughput of signalized arterials," in *IEEE Intelligent Vehicles Symposium*, 2017.

[9] R. E. Stern, S. Cui, M. L. Delle Monache, R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, H. Pohlmann, F. Wu, B. Piccoli *et al.*, "Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments," *Transportation Research Part C: Emerging Technologies*, 2018.

[10] C. Wu, A. M. Bayen, and A. Mehta, "Stabilizing traffic with autonomous vehicles," in *International Conference on Robotics and Automation*, 2018.

[11] L. Jin, M. Čičič, S. Amin, and K. H. Johansson, "Modeling the impact of vehicle platooning on highway congestion: A fluid queuing approach," in *Conference on Hybrid Systems: Computation and Control*. ACM, 2018.

[12] S. Sivaranjani, Y.-S. Wang, V. Gupta, and K. Savla, "Localization of disturbances in transportation systems," in *IEEE Conference on Decision and Control (CDC)*, 2015.

[13] D. A. Lazar, E. Bıyık, D. Sadigh, and R. Pedarsani, "Learning how to dynamically route autonomous vehicles on shared roads," *arXiv preprint arXiv:1909.03664*, 2019.

[14] M. Salazar, M. Tsao, I. Aguiar, M. Schiffer, and M. Pavone, "A congestion-aware routing scheme for autonomous mobility-on-demand systems," in *European Control Conference (ECC)*, 2019.

[15] T. Roughgarden and É. Tardos, "How bad is selfish routing?" *Journal of the ACM (JACM)*, 2002.

[16] L. Fleischer, K. Jain, and M. Mahdian, "Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games," in *Symposium on Foundations of Computer Science*. IEEE, 2004.

[17] P. N. Brown and J. R. Marden, "The robustness of marginal-cost taxes in affine congestion games," *IEEE Transactions on Automatic Control*, 2017.

[18] W. H. Sandholm, "Evolutionary implementation and congestion pricing," *The Review of Economic Studies*, 2002.

[19] N. D. Daw, J. P. O'doherty, P. Dayan, B. Seymour, and R. J. Dolan, "Cortical substrates for exploratory decisions in humans," *Nature*, 2006.

[20] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, 1994.

[21] P. Koster, E. Verhoef, S. Shepherd, and D. Watling, "Preference heterogeneity and congestion pricing: The two route case revisited," *Transportation Research Part B: Methodological*, vol. 117, pp. 137–157, 2018.

[22] A. Sumalee and W. Xu, "First-best marginal cost toll for a traffic network with stochastic demand," *Transportation Research Part B: Methodological*, vol. 45, no. 1, pp. 41–59, 2011.

[23] W. Krichene, J. D. Reilly, S. Amin, and A. M. Bayen, "Stackelberg routing on parallel transportation networks," *Handbook of Dynamic Game Theory*, 2018.

[24] A. Y. Ng, S. J. Russell *et al.*, "Algorithms for inverse reinforcement learning." in *International Conference on Machine Learning (ICML)*, 2000.

[25] P. Abbeel and A. Y. Ng, "Exploration and apprenticeship learning in reinforcement learning," in *International Conference on Machine Learning (ICML)*. ACM, 2005.

[26] B. D. Ziebart, A. L. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning." in *AAAI*, 2008.

[27] W. Cheng, J. Fürnkranz, E. Hüllermeier, and S.-H. Park, "Preference-based policy iteration: Leveraging preference learning for reinforcement learning," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2011, pp. 312–327.

[28] N. Kallus and M. Udell, "Revealed preference at scale: Learning personalized preferences from assortment choices," in *Proceedings of the 2016 ACM Conference on Economics and Computation*, 2016, pp. 821–837.

[29] M. Zadimoghaddam and A. Roth, "Efficiently learning from revealed preference," in *International Workshop on Internet and Network Economics*. Springer, 2012, pp. 114–127.

[30] E. Bıyık, D. A. Lazar, R. Pedarsani, and D. Sadigh, "Altruistic autonomy: Beating congestion on shared roads," in *Workshop on the Algorithmic Foundations of Robotics*, 2018.

[31] C. C. Gonzaga, "Path-following methods for linear programming," *SIAM review*, 1992.

[32] C. K. Chau and K. M. Sim, "The price of anarchy for non-atomic congestion games with symmetric cost maps and elastic demands," *Operations Research Letters*, 2003.

[33] R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban, "An interior algorithm for nonlinear optimization that combines line search and trust region steps," *Mathematical programming*, 2006.

[34] G. Andrew and J. Gao, "Scalable training of l 1-regularized log-linear models," in *International Conference on Machine Learning (ICML)*. ACM, 2007.

[35] E. Biyik, N. Huynh, M. J. Kochenderfer, and D. Sadigh, "Active preference-based gaussian process regression for reward learning," in *Proceedings of Robotics: Science and Systems (RSS)*, July 2020.

[36] E. Biyik, M. Palan, N. C. Landolfi, D. P. Losey, and D. Sadigh, "Asking easy questions: A user-friendly approach to active reward learning," in *Conference on Robot Learning (CoRL)*, 2019.

[37] P. N. Brown and J. R. Marden, "Fundamental limits of locally-computed incentives in network routing," in *IEEE American Control Conference (ACC)*, 2017.

[38] R. D. Luce, *Individual choice behavior: A theoretical analysis*. Courier Corporation, 2012.

VII. APPENDIX

Proof of Lemma 1. We begin by noting that for any given network and feasible flow demand, a *continuum* of equilibria exist which satisfy the flow demand, where different equilibria have positive flow on different sets of roads. This fact stems from the two regimes (free-flow and congested) that exist, as well as the two different vehicle types. We note that for a congested road, a lower density yields a higher vehicle flow and lower latency. Accordingly, via the relationship with density, the latency function decreases smoothly with an increase in the flow of regular or autonomous cars. Note this does not mean that transferring flow to a road in a dynamic setting decreases its latency, rather that the function dictating the relationship between latency and flow in the congested regime is smoothly decreasing (in accordance with each quantity's relationship with the corresponding vehicle density).

With this in mind, we consider the continuum of equilibria that exist. Consider a Nash Equilibrium with positive flow on roads $[m]$. Assume road m is congested, as otherwise the lemma would be satisfied with $m' = m$. Roads $[m-1]$ must be congested so as to have the same latency as road m, as it is an equilibrium. We constructively find a different equilibrium as follows. Consider another configuration serving the same flow demand, where the flow on road m serves less flow and roads $[m-1]$ serve more flow as follows. Note that in each configuration serving the flow demand, vehicle flow is conserved but vehicle density is not. Accordingly, road m has higher latency (from higher density) and roads $[m-1]$ have lower latency (from lower density). Since the flow-latency relationship in (3) has latency as a monotonically decreasing function with flow of each vehicle type, we can consider a new configuration where each road in $[m-1]$ has equal latency (which is less than the first equilibrium configuration) while serving more flow. Accordingly, we can construct one of the two following configurations.

1) the latencies on roads $[m-1]$ are reduced to a_m, the free-flow latency on road m, or
2) road m serves no flow and the latencies on roads $[m-1]$ are greater than a_m.

In the first case, the lemma is satisfied with $m' = m$. In the second case, we can consider the same logic again, again considering a new configuration with less flow on road $m-1$ and more flow on roads $[m-2]$. This continues until either we achieve case 1 above or until we are reduced to a single road. If that occurs, traffic can be routed in free-flow on that road, since any feasible flow in the congested regime is less than the maximum free-flow on a road. $\qquad\square$

Proof of Theorem 1. The definition of Nash Equilibrium and the fact that latency on a road is always equal to or greater than its free-flow latency together imply that at Nash Equilibrium, if road m has positive flow then all roads with free flow latency less than a_m have positive flow as well. These also imply that if a road m in free-flow has positive flow, roads with greater free-flow latency will have zero flow. Further, we use Lemma 1 to show that all routings in the set of BNE will have one road in free-flow. Assume for the purposes of contradiction that we have a routing in the set of BNE in which only roads

$[m]$ have positive flow and all are congested, with equilibrium latency $\hat{\ell}_0$. The total cost is then $\hat{\ell}_0(\lambda^{\mathrm{h}} + \lambda^{\mathrm{a}})$, where $\hat{\ell}_0 > a_m$. By Lemma 1, another routing exists in the set of NE which uses roads $[m']$, where $m' \leq m$ and road m' is in free flow. The cost of this equilibrium is $a_{m'}(\lambda^{\mathrm{h}} + \lambda^{\mathrm{a}}) \leq a_m(\lambda^{\mathrm{h}} + \lambda^{\mathrm{a}}) < \hat{\ell}_0(\lambda^{\mathrm{h}} + \lambda^{\mathrm{a}})$ contradicting our premise.

So far we have proved the numbered claims. We prove the remaining claim by contradiction. Assume there are two routings in f, $f' \in \mathrm{BNE}(\lambda^{\mathrm{h}}, \lambda^{\mathrm{a}})$ which have different free-flow roads, m and m' respectively. Assumption 1 implies $a_m \neq a_{m'}$; let $a_m < a_{m'}$. Since all selfish users experience the same latency, the total latency of routing f is $(\lambda^{\mathrm{h}} + \lambda^{\mathrm{a}})a_m < (\lambda^{\mathrm{h}} + \lambda^{\mathrm{a}})a_{m'}$, which is the total latency of routing f'. However by the definition of BNE, the total latency of the two routings are equal, yielding a contradiction. $\qquad\square$

Proof of Theorem 2. The first property directly follows from Theorem 1, as the regular vehicles have to be at a Nash equilibrium due to selfishness. To prove the second property, we note that for roads that have higher latencies than road m_{EQ}^*, a Nash equilibrium is not necessary due to flexibility. As $\ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i)$ is a non-increasing continuous function of f_i^{a} and decreasing for $s_i = 1$, roads that have higher latencies than road m_{EQ}^* will always be in free-flow.

Now we assume some of the roads with indices greater than m_{EQ}^* and less than m_{ALL}^* are in free-flow, but not at maximum flow. Then we could simply transfer some flow from the road m_{ALL}^* to those roads and have lower overall costs. This is a contradiction, completing the proof for the third property. $\qquad\square$

Proof of Theorem 3. Since road m_{EQ}, the longest equilibrium road, is in free-flow by Theorem 1 and m_{EQ}^* is the minimum feasible m_{EQ}, our solution is restricted to the set of BNE. Accordingly, we can restrict our optimization to routing in which the longest equilibrium road is in free-flow. This allows us to write an optimization equivalent to checking the feasibility of a routing with the desired congestion profile:

$$\max_{\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}} \in \mathbb{R}_{\geq 0}^N} 1$$
$$\text{s.t.} \sum_{i \in [m_{\mathrm{EQ}}^*]} f_i^{\mathrm{h}} = \lambda^{\mathrm{h}}, \quad \sum_{i \in [m_{\mathrm{EQ}}^*]} f_i^{\mathrm{a}} = \lambda^{\mathrm{a}}$$
$$\ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, 1) = a_{m_{\mathrm{EQ}}^*} \; \forall i \in [m_{\mathrm{EQ}}^* - 1]$$
$$f_{m_{\mathrm{EQ}}^*}^{\mathrm{h}} + f_{m_{\mathrm{EQ}}^*}^{\mathrm{a}} \leq \bar{F}_{m_{\mathrm{EQ}}^*}(f_{m_{\mathrm{EQ}}^*}^{\mathrm{h}}, f_{m_{\mathrm{EQ}}^*}^{\mathrm{a}})$$

The constraints can be shown to be affine in the decision variables. As it is a linear program, it can be solved in $O(N^3)$ time [31]. Finding m_{EQ}^* requires a search in $O(N)$ time. $\qquad\square$

Proof of Theorem 5. We first prove, similar to Luce's choice axiom [38], changing the latency or the price of some roads does not alter the autonomous flow ratio between the other options, including the alternative option. For this, we look at $\mathbb{E}_{\boldsymbol{\omega}, \zeta}\left[\frac{P(i_1 \mid \boldsymbol{\omega}, \zeta, \boldsymbol{\ell}, \boldsymbol{p})}{P(i_2 \mid \boldsymbol{\omega}, \zeta, \boldsymbol{\ell}, \boldsymbol{p})}\right]$, where i_2 is an undominated option, $P(i \mid \boldsymbol{\omega}, \zeta, \boldsymbol{\ell}, \boldsymbol{p})$ is the probability of choosing option $i \in [N] \cup \{0\}$ under the given reward parameters, latency and price. We note

$$P(i \mid \boldsymbol{\omega}, \zeta, \boldsymbol{\ell}, \boldsymbol{p}) = \frac{\exp(r(\boldsymbol{\ell}, \boldsymbol{p}, i; \boldsymbol{\omega}, \zeta))}{\sum_{i'=0}^N \exp(r(\boldsymbol{\ell}, \boldsymbol{p}, i'; \boldsymbol{\omega}, \zeta))}, \quad \text{so}$$
$$\mathbb{E}_{\boldsymbol{\omega}, \zeta}\left[\frac{P(i_1 \mid \boldsymbol{\omega}, \zeta, \boldsymbol{\ell}, \boldsymbol{p})}{P(i_2 \mid \boldsymbol{\omega}, \zeta, \boldsymbol{\ell}, \boldsymbol{p})}\right] = \mathbb{E}_{\boldsymbol{\omega}, \zeta}\left[\frac{\exp(r(\boldsymbol{\ell}, \boldsymbol{p}, i_1; \boldsymbol{\omega}, \zeta))}{\exp(r(\boldsymbol{\ell}, \boldsymbol{p}, i_2; \boldsymbol{\omega}, \zeta))}\right].$$

As the reward of an option depends only on that option's price and latency, this proves the first statement above.

Equipped with this result, we now prove Theorem 5. Assume the optimal solution is such that all human-driven flow is in congested roads. Let k be the index of the highest free-flow latency road with nonzero human-driven flow. We first show there exists an equally optimal solution with no dominated roads in $[k]$. For this, we simply set the prices of roads $[k]$ such that they are all equal and the total autonomous demand in $[k]$ in the original and the new solution is the same. Since the new solution has no dominated roads in $[k]$, their new price has to be at least as high as the undominated roads in $[k]$ of the original solution, which implies the profit constraint is still satisfied. As the total demand served and the overall latency values are the same between these two solutions by the first statement above, the two solutions are equally good. In the remaining of the proof, we refer to the new solution as the "optimal solution" for clarity, and show there exists a better solution, leading to a contradiction.

Denote the optimal solution with $(\boldsymbol{\ell}^*, \boldsymbol{p}^*, \boldsymbol{f}^{\mathrm{h}*}, \boldsymbol{f}^{\mathrm{a}*})$, and its dominated options with D^*, noting $[k] \cap D^* = \emptyset$. Let the ratio of autonomous service users in road i to the autonomous service users who decline the service be β_i, i.e., $\beta_i = \frac{f^{\mathrm{a}*}_i}{\lambda^{\mathrm{a}} - \sum_{i' \in [N]} f^{\mathrm{a}*}_{i'}}$.

Let $k' = \arg\max_i a_i$ subject to $i \leq k$ and $\left(\sum_{i' \in [k]} f^{\mathrm{h}*}_{i'}, \sum_{i' \in [k]} f^{\mathrm{a}*}_{i'} \right)$ can be allocated into $[k']$ when all roads in $[k']$ have latency equal to $a_{k'}$. Existence of such a k' is guaranteed by Lemma 1.

Using k', we propose an alternative solution $(\boldsymbol{\ell}', \boldsymbol{p}', \boldsymbol{f}^{\mathrm{h}'}, \boldsymbol{f}^{\mathrm{a}'})$:

$$
\ell'_i = \begin{cases} \ell^*_i & \text{if } i \in [N] \setminus [k] \\ a_i & \text{if } i \in [k] \setminus [k'] \\ a_{k'} & \text{if } i \in [k'] \end{cases},
$$

$$
p'_i = \begin{cases} p^*_i & \text{if } i \in [N] \setminus ([k] \cup D^*) \\ p^*_i + \epsilon & \text{if } i \in [k] \cup D^* \end{cases}.
$$

where $\epsilon \geq 0$ is such that the ratio of autonomous service users in $[k']$ to the autonomous service users who decline the service in the alternative solution is equal to $\sum_{i \in [k]} \beta_i$. Since $\ell'_i < \ell^*_i$ for $\forall i \in [k']$, $\omega_2 > 0$ for all users, and $[k] \cap D^* = \emptyset$, the existence of such an ϵ is guaranteed.

By the first statement, we know the ratios (β_i's) in the alternative solution will be equal to the optimal solution for roads in $[N] \setminus ([k] \cup D^*)$ as their latencies and prices are the same. Similarly, roads in D^* are also dominated in the alternative solution, because no road has higher latency and prices went up maximally (by ϵ) for those roads. This means the ratios are equal between two solutions for all roads in $[N] \setminus [k]$. Noting roads in $[k] \setminus [k']$ are dominated in the alternative solution, as they have the same price as the roads in $[k']$ but higher latency, we write the total autonomous demand:

$$
\lambda^{\mathrm{a}} = \left(\lambda^{\mathrm{a}} - \sum_{i \in [N]} f^{\mathrm{a}'}_i \right) (1 + \beta_1 + \beta_2 + \cdots + \beta_N)
$$

by the construction of ϵ. As the total autonomous demand is the same between two solutions, we get $\lambda^{\mathrm{a}} - \sum_{i \in [N]} f^{\mathrm{a}'}_i = \lambda^{\mathrm{a}} - \sum_{i \in [N]} f^{\mathrm{a}*}_i$ meaning the flow of users who decline the

service is the same. This implies $f^{\mathrm{a}'}_i = f^{\mathrm{a}*}_i$ for $\forall i \in [N]/[k]$, and $\sum_{i \in [k]} f^{\mathrm{a}'}_i = \sum_{i \in [k]} f^{\mathrm{a}*}_i$. As the selection of k' ensures there is enough room for selfish vehicles in $[k']$ in the alternative solution, we have $\sum_{i \in [k']} f^{\mathrm{h}'}_i = \sum_{i \in [k]} f^{\mathrm{h}*}_i$.

Finally, the alternative solution satisfies the profit constraint because the price and the autonomous flow in $[N] \setminus [k]$ are the same as the optimal solution, and the remaining autonomous flow pays higher price in the alternative solution.

Overall, the alternative solution is feasible, serves the same amount of flow, but has lower latency overall, $(\boldsymbol{\ell}^*, \boldsymbol{p}^*, \boldsymbol{f}^{\mathrm{h}*}, \boldsymbol{f}^{\mathrm{a}*})$ cannot be the optimal solution. $\square$

Role of Assumption 1. Without this assumption, m^*_{EQ} in Theorem 1 would instead represent a set of roads, all of which would be in free-flow. Theorem 2 would change similarly to Theorem 1. Theorem 3 is not altered, the constraint associated with m^*_{EQ} would instead apply to all roads within the set. Finally, Theorem 4 would remain the same as well. What would potentially change this theorem is if the set of roads in m^*_{EQ} could have different congestion levels in the computed BFNE. To see why this is not the case, note that the flexibility profile is with respect to the quickest route available to users (*i.e.* the least congested road in the set m^*_{EQ}). Accordingly, having another road in the set be more congested would not help satisfy any constraint related to the flexibility profile and would also not serve more flow than if it had the same latency as the minimum latency road in m^*_{EQ}. Hence, all roads in m^*_{EQ} will have the same latency, so the computational complexity remains the same. Theorem 5 would still hold: there exists a free-flow road used by human drivers in the optimal solution.

Erdem Bıyık is a fourth-year Ph.D. candidate in the EE Department at Stanford. He received his B.Sc. degree in electrical and electronics engineering from Bilkent University, Turkey, in 2017. He is interested in enabling robots to learn from various forms of human feedback, and designing robot policies to improve the performance of multi-agent systems both in cooperative and competitive settings.

Daniel A. Lazar is a fifth-year Ph.D. candidate in the ECE Department at the University of California, Santa Barbara. In 2014, he received a B.Sc. in Electrical Engineering from Washington University in St. Louis, after which he spent two years working in wireless communications. His research focuses on control of transportation networks.

Ramtin Pedarsani is an Assistant Professor in the ECE Department at the University of California, Santa Barbara. He received the B.Sc. degree in electrical engineering from the University of Tehran in 2009, the M.Sc. degree in communication systems from the Swiss Federal Institute of Technology (EPFL) in 2011, and his Ph.D. from the University of California, Berkeley, in 2015. His research interests include networks, game theory, machine learning, and transportation systems.
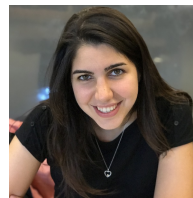
Dorsa Sadigh is an Assistant Professor in the CS and EE departments at Stanford University. Her research interests lie in the intersection of robotics, learning and control theory. Specifically, she is interested in developing algorithms for safe and adaptive human-robot interaction. Dorsa has received her doctoral degree in Electrical Engineering and Computer Sciences (EECS) at UC Berkeley in 2017, and has received her B.Sc. in EECS at UC Berkeley in 2012.