Value of Information in Bayesian Routing Games

Manxi Wu, Saurabh Amin, and Asuman E. Ozdaglar *

Abstract

We study a routing game in an environment with multiple heterogeneous information systems and an uncertain state that affects edge costs of a congested network. Each information system sends a noisy signal about the state to its subscribed traveler population. Travelers make route choices based on their private beliefs about the state and other populations’ signals. The question then arises, “How does the presence of asymmetric and incomplete information affect the travelers’ equilibrium route choices and costs?” We develop a systematic approach to characterize the equilibrium structure, and determine the effect of population sizes on the relative value of information (i.e., difference in expected traveler costs) between any two populations. This effect can be evaluated using a population-specific size threshold. One population enjoys a strictly positive value of information in comparison to the other if and only if its size is below the corresponding threshold. We also consider the situation when travelers may choose an information system based on its value, and characterize the set of equilibrium adoption rates delineating the sizes of subscribed traveler populations. The resulting routing strategies are such that all travelers face an identical expected cost and no traveler has the incentive to change her subscription.

Index Term: Games/group decisions: Nonatomic; Transportation: Technology; Information systems.

1 Introduction

Travelers are increasingly relying on traffic navigation services to make their route choice decisions. In the past decade, numerous services have come to the forefront, including Waze/Google maps, Apple maps, INRIX, etc. These Traffic Information Systems (TISs) provide their subscribers with costless information about the uncertain network condition (state), which is typically influenced by exogenous factors such as weather, incidents, and road conditions. The information provided by TIS can be especially useful in making travel decisions when a change in state corresponds to changes in travel times of multiple
edges or routes of the network. Experiential evidence suggests that the accuracy levels of TISs are less than perfect, and exhibit heterogeneities due to the inherent technological differences in data collection and analysis procedures. Moreover, travelers may use different TISs or choose not to use them at all, depending on factors such as marketing, usability, and availability. Therefore, we can reasonably expect that travelers face an environment of asymmetric and incomplete information about the network state.

Importantly, information heterogeneity can directly influence the travelers' route choice decisions, and the resulting congestion externalities. Consider an example where some travelers are informed by their TIS that a certain route has an incident. Taking a detour based on this information may not only reduce their own travel time, but also benefit the uninformed travelers by shifting traffic away from the affected route. However, if too many travelers take the detour, then this alternate route will also start getting congested, limiting the benefits of information (Section 2). Thus, the question arises as to how information heterogeneity impacts the travelers' route choices and costs.

In this article, we develop a game-theoretic approach to study this question. We consider a routing game in which the travelers are privately informed about the network state by their respective TIS, and choose strategies based on their beliefs about the state and other travelers' behavior. We conduct a systematic study of how the sizes of heterogeneously informed traveler populations affect the equilibrium structure and the relative value of information faced by travelers subscribed to one TIS in comparison to another TIS. Furthermore, we characterize the set of equilibrium adoption rates when travelers can choose any available TIS based on its relative value.

1.1 Our Model and Contributions

We model the traffic routing problem in an asymmetric and incomplete information environment as a Bayesian routing game. We consider a general traffic network with an uncertain state that is realized from a finite set according to a prior probability distribution. The cost function (travel time) of each edge in the network is increasing in the aggregate traffic load on that edge. Moreover, the edge costs are state-dependent in that the state can affect them in various ways. There are multiple heterogeneous TISs, each sending a noisy signal of the state to its subscribed traveler population. The signal sent by each TIS is privately known to only its subscribers, but the joint distribution of the state and all signals is known by all travelers (common knowledge). This joint distribution is the common prior of the game, and each population's private belief of the state and other populations' private signals is derived from it. Our information environment is general in that we do not impose any structural assumption on the common prior. For example, the signals of different TISs can be correlated or independent, conditional on the state. Also, one TIS can be more accurate than another in some states, but less in the other states. Thus, we do not assume that the TISs are ordered according to the accuracy of their signals (Section 3).

We use Bayesian Wardrop Equilibrium (BWE) as the solution concept of our game. In a BWE, all populations assign demand on routes with the smallest expected cost based on their private beliefs. In fact, our game is a weighted potential game, i.e. the set of
BWE is the optimal solution set of a convex optimization problem that minimizes the weighted potential function over the set of feasible routing strategies of traveler populations (Theorem 1). This property establishes the essential uniqueness of BWE, i.e. the equilibrium edge load vector is unique. However, the strategy-based optimization problem is not directly useful for analyzing how the set of equilibrium strategies and population costs depend on the population size vector. To address this issue, we provide a characterization of the set of feasible route flows (i.e. route flows induced by feasible strategy profiles) as a convex polytope (Proposition 1), and show that this set is the optimal solution set of another convex optimization problem (Proposition 2). The optimal value of this flow-based formulation is the value of weighted potential function in equilibrium (Section 4).

The flow-based formulation enables us to analyze the sensitivity of equilibrium structure with respect to perturbations in the size vector. The constraints in this formulation include: basic route feasibility constraints, a set of size-independent equality constraints, and a set of size-dependent inequality constraints. The equality constraints ensure that any shift in route flows resulting from a change of signal received by one population is independent of the signals received by other populations. Each inequality constraint corresponds to a single population, and ensures that the maximum extent to which the received signals impact its equilibrium routing behavior is limited by the population’s size. Hence, we refer to these inequality constraints as information impact constraints (IICs). Consequently, the effects of perturbation in the size vector on the equilibrium structure can be studied by evaluating the tightness of the IICs corresponding to the perturbed populations at the optimum of the flow-based formulation.

In particular, Theorem 2 describes how the qualitative properties of equilibrium route flows change under perturbations in the sizes of any two populations, with sizes of all other populations being fixed (i.e. directional perturbations of the size vector). Among the two perturbed populations, we say that one population is the “minor population” if its size is smaller than a certain (population-specific) threshold. Then, the corresponding IIC is tight in equilibrium, i.e. the impact of information on the minor population is fully attained. These population-specific thresholds depend on the common prior distribution as well as sizes of all other populations. In fact, each threshold can be computed by solving a linear program. Based on the two thresholds, we can distinguish three qualitatively distinct equilibrium regimes: In the two side regimes exactly one population assumes the minority role, on the other hand, in the middle regime neither population is minor.

We can apply Theorem 2 to analyze the sensitivity of the equilibrium value of the weighted potential function under directional perturbation of population sizes (Proposition 3). In the middle regime, perturbing the relative sizes does not change the equilibrium value of the potential function. On the other hand, in the two side regimes the value of the potential function monotonically decreases as the size of the minor population increases. Thanks to the essential uniqueness of BWE, the equilibrium edge load vector does not change with this directional perturbation if and only if the size vector falls in the middle regime.

These results allow us to compare the expected cost in equilibrium faced by travelers in any pair of populations. In particular, we can evaluate how this cost difference —
which we call the relative value of information — changes with pairwise size perturbations. By using the results on sensitivity analysis of general convex optimization problems (Fiacco et al. 1983, Rockafellar 1984), we show that the relative value of information is proportional to the derivative of the equilibrium value of potential function in the direction of perturbation. Combining this observation with Proposition 3 we obtain that the minor population faces a lower cost relative to the other population in the two side regimes; whereas both populations face identical costs in the middle regime. Importantly, the relative value of information is non-increasing in the size of its subscribed population (Theorem 3). This result is based on the intuition that an individual traveler faces higher congestion externality when more travelers have access to the same information, and hence make their route choices according to the same strategy. Thus, an increase in the size of minor population decreases the relative imbalance in congestion externality, thereby reducing the advantage enjoyed by its travelers over the other population.

Our results can be easily specialized to a simpler information environment in which one of the populations does not have an access to TIS (uninformed population). In this case, we obtain that the equilibrium cost of the uninformed population is no less than that of any other population regardless of the size vector (Proposition 4). That is, having access to a TIS always leads to a non-negative relative value of information in comparison to being uninformed (Section 5).

Furthermore, we extend our approach of comparing the relative value of information for any pair of populations to study how the equilibrium outcome depends on the population sizes in general. In particular, we characterize a non-empty set of size vectors, for which the equilibrium edge load vector is size-independent, and the value of potential function is minimized (Proposition 5). This result enables us to study situations when travelers can choose their TIS subscription based on its value. Then, the notion of adoption rate of TIS becomes relevant, because the value of information to users (as well as non-users) of a TIS depends on the size of the population with access to information provided by it. Intuitively, if it is possible for the non-users of a particular TIS to receive a greater value by adopting it, then they may continue to do so until there is no longer a positive relative value in adopting that TIS. Theorem 4 shows that the set of equilibrium adoption rates (i.e. fraction of travelers choosing each TIS in equilibrium) is exactly the set of population sizes that minimize the value of potential function at equilibrium. Naturally, any size vector that is induced by travelers’ equilibrium choice of TIS satisfies the property that all travelers face identical costs and no traveler has an incentive to change her TIS subscription (Section 6).

We conclude our work in Section 7. We include supplementary results, examples, and most of the proofs in the electronic companion of this article.

### 1.2 Related Work

**Congestion games.** Well-known results in classical congestion games include their equivalence with potential games (Rosenthal 1973, Monderer and Shapley 1996, and Sandholm 2001), analysis of network formation games as congestion games (Gopalakrishnan et al. 2014, and Tardos and Wexler 2007), and equilibrium inefficiency (Roughgarden and Tardos 2007).
Koutsoupias and Papadimitriou [1999], Correa et al. [2007], Acemoglu and Ozdaglar [2007], and Nikolova and Stier-Moses [2014]). Milchtaich [1996] has studied congestion games with player-specific cost functions, which can model both heterogeneous private information and heterogeneous preferences. However, the existence of potential function is not guaranteed with general player-specific cost functions. In our game, the heterogeneity in the expected costs among populations arises only due to heterogeneous private beliefs, which are derived from a common prior. This feature makes our game a weighted potential game (Sandholm [2001]). Moreover, Acemoglu et al. [2018] studied a congestion game where travelers have different information sets about the available edges (routes). While their work focuses on heterogeneous information about the network structure, we use a Bayesian approach to model the information heterogeneity resulting from the differences in TIS access and accuracy.

Traffic Information Systems. Prior work has studied the effects of TIS on travelers’ departure time choices (Arnott et al. [1991], and Khan and Amin [2018]), and on their route choices (Ben-Akiva et al. [1991], Ben-Akiva et al. [1996]). In particular, the paper Mahmassani and Jayakrishnan [1991] used simulation approach to show that as more travelers receive information, the informed travelers gradually start facing higher costs and their relative value of being informed diminishes. Our analysis, when applied to the game with one uninformed population and other more informed ones, also leads to similar conclusion. Our results are more general because they are applicable to routing games with multiple heterogeneously informed populations with arbitrary TIS accuracies.

Value of Information. In a classical paper, Blackwell [1953] showed that for a single decision maker, more informative signal always results in higher expected utility. In game-theoretic settings, it is generally difficult to determine whether the value of information in equilibrium for individual players and/or society is positive, zero or negative (see Hirshleifer [1971], and Haenfler [2002]). However, the value of information is guaranteed to be positive when certain conditions are satisfied; see for example Neyman [1991], and Lehrer and Rosenberg [2006]. Since travelers are non-atomic players in our game, the relative value of information between any two TISs is equivalent to the value of information for an individual traveler when her subscription changes unilaterally. We give precise conditions on the population sizes under which the value of information in our Bayesian routing game is positive, zero, or negative.

2 Motivating Example

In this section, we motivate our analysis using a simple game of two asymmetrically informed traveler populations routing over a network of two parallel routes, denoted $r_1$ and $r_2$. The network state $s$ belongs to the set $S = \{a, n\}$, where the state $a$ represents an incident condition on $r_1$, and the state $n$ represents the nominal condition. The state $a$ occurs with probability $p \in (0, 1)$. The network faces a unit size of demand ($D = 1$), which is comprised of two traveler populations: population 1 with size $\lambda^1$ and population 2 with size $\lambda^2 = 1 - \lambda^1$. Each population $i \in I = \{1, 2\}$ receives a signal $t^i \in T^i = \{a, n\}$ from its TIS. Assume for simplicity that population 1 receives the correct state with probability 1.
(i.e. complete information), and population 2 receives signal \( a \) or \( n \) with probability 0.5, independent of the state (i.e. no information).

Let \( q_i^1(t^i) \) denote the traffic demand assigned to route \( r_1 \) by population \( i \) when receiving signal \( t^i \); the remaining demand \( (\lambda^i - q_i^1(t^i)) \) is assigned to route \( r_2 \). Since the signal \( t^2 \) is independent of the state, we have \( q_i^2(a) = q_i^2(n) = q_i^2 \). A feasible demand assignment must satisfy the constraints: \( 0 \leq q_i^1(t^i) \leq \lambda^i \) and \( 0 \leq q_i^2 \leq \lambda^2 \). For this example, we can represent a routing strategy profile as \( q_1 = (q_1^1(a), q_1^1(n), q_1^2) \). We denote the aggregate route flow on \( r \) as \( f_r \).

We consider simple affine cost functions: the cost of \( r_1 \) is \( c^a_1(f_1) = \alpha_1^a f_1 + b \) in state \( a \), and \( c^a_1(f_1) = \alpha_1^n f_1 + b \) in state \( n \); the cost of \( r_2 \) is \( c_2(f_2) = \alpha_2 f_2 + b \). The slopes satisfy \( \alpha_1^a < \alpha_2 < \alpha_1^n \). Since population 1 has complete information, its travelers know the exact cost function in both states. However, since population 2 travelers are uninformed, they make their route choices based on the expected cost of each route, evaluated according to the prior distribution of states.

This routing game with heterogeneously informed traveler populations admits a Bayesian Wardrop equilibrium, as discussed in Section 3. Let \( q_i^* = (q_i^{1*}(a), q_i^{1*}(n), q_i^{2*}) \) denote an equilibrium strategy profile. Each population, given the signal it receives, can either assign all its demand on one of the two routes, or splits on both routes. Thus, there are \( 3^3 = 27 \) possible cases. Our results can be used to study how the equilibrium strategies and route flows change as the size of a population varies from 0 to 1. Detailed analysis for this simple routing game and some interesting variants are available in [Wu et al. 2017]. Specifically, we find that there exists a threshold size of population 1, \( 0 < \Delta^1 = \alpha_2 \left( \frac{1}{\alpha_1^a + \alpha_2} - \frac{1}{\alpha_1^n + \alpha_2} \right) < 1 \), such that the qualitative structure of equilibrium routing strategies is different based on whether \( \lambda^1 \in [0, \Delta^1) \) or \( \lambda^1 \in [\Delta^1, 1] \).

In the first regime, i.e. when \( \lambda^1 \in [0, \Delta^1) \), the game admits a unique equilibrium: \( q_1^{1*}(a) = 0, q_1^{1*}(n) = \lambda^1, \) and \( q_1^{2*} = \frac{\alpha_2}{\alpha_1^a + \alpha_2} - \lambda^1(1-p)(\alpha_1^n + \alpha_2) \). This equilibrium regime corresponds to the following outcome: in state \( n \) (resp. state \( a \)), population 1 assigns all its demand on route \( r_1 \) (resp. route \( r_2 \)), and population 2 splits its demand on both routes. The induced equilibrium flow on \( r_1 \) is given by \( f_1^* = q_1^{2*} \) if \( t^1 = a \), and \( f_1^* = \lambda^1 + q_1^{2*} \) if \( t^1 = n \).

On the other hand, in the second regime, i.e. when \( \lambda^1 \in [\Delta^1, 1] \), the equilibrium set may not be singleton, and can be represented as follows: \( q_1^{1*}(a) = \chi, q_1^{1*}(n) = \Delta^1 + \chi, \) and \( q_1^{2*} = \frac{\alpha_2}{\alpha_1^n + \alpha_2} - \chi \), where \( \max \left\{ 0, \lambda^1 - \frac{\alpha_1^a}{\alpha_1^a + \alpha_2} \right\} \leq \chi \leq \min \left\{ \frac{\alpha_2}{\alpha_1^n + \alpha_2}, \lambda^1 - \Delta^1 \right\} \). Thus, the equilibrium set is a one-dimensional interval for \( \lambda^1 \in (\Delta^1, 1) \), and a singleton set for \( \lambda^1 = \Delta^1 \) or \( \lambda^1 = 1 \). In this regime, both populations face identical expected route costs in equilibrium. Consequently, each population splits its demand on both routes. Moreover, the equilibrium route flow on each route is unique and independent of \( \lambda^1 \): \( f_1^* = \frac{\alpha_2}{\alpha_1^n + \alpha_2} \) if \( t^1 = a \), and \( f_1^* = \frac{\alpha_2}{\alpha_1^a + \alpha_2} \) if \( t^1 = n \).

Notice that when \( \lambda^1 \in [0, \Delta^1) \), we have \( q_1^{1*}(n) - q_1^{1*}(a) = \lambda^1 \), i.e. population 1 shifts all its demand to \( r_2 \) when receiving the signal about the incident on \( r_1 \). However, if \( \lambda^1 \in [\Delta^1, 1] \), we have \( q_1^{1*}(n) - q_1^{1*}(a) = \Delta^1 < \lambda^1 \), i.e. the change in the received signal only influences a part of travelers in population 1. One can say that the information impacts the entire demand of population 1 in the first regime, but not in the second regime.
For any feasible $\lambda = (\lambda^1, \lambda^2)$, we can calculate the equilibrium population costs, denoted $C^i_*(\lambda)$. If $\lambda^1 \in [0, A^1)$, it is easy to check that $C^{2*}(\lambda) - C^{1*}(\lambda) > 0$, i.e. when the state information is only available to a small fraction of travelers, the informed travelers have an advantage over the uninformed ones. On the other hand, if the size of informed population exceeds the threshold $A^1$, then $C^{1*}(\lambda) = C^{2*}(\lambda)$. In this case, all travelers face identical cost in equilibrium.

We illustrate the aforementioned results in Fig. 1 using the following parameters: $\alpha^1_1 = 1, \alpha^1_2 = 3, \alpha_2 = 2, b = 20$, and $p = 0.2$. The costs are normalized by the socially optimal cost, denoted $C^{so}$, which is the minimum cost achievable by a social planner with complete information of the state. Our subsequent analysis considers more general settings with multiple (heterogeneous and possibly correlated) information systems and general network topology.

Figure 1: Effects of varying population 1 size on equilibrium structure and costs: (a) Population strategies, (b) Flow on route $r_1$, (c) Population costs.

3 Model

3.1 Environment

To generalize the simple routing game in Section 2, we consider a transportation network modeled as a directed graph. For ease of exposition, we assume that the network has a single origin-destination pair. All our results apply to networks with multiple origin-destination pairs; see Sec 4. Let $\mathcal{E}$ denote the set of edges and $\mathcal{R}$ denote the set of routes. The finite set of network states, denoted $\mathcal{S}$, represents the set of possible network conditions, such as incidents, weather, etc. The network state, denoted $s$, is randomly drawn by a fictitious player “Nature” from $\mathcal{S}$ according to a distribution $\theta \in \Delta(\mathcal{S})$, which determines the prior probability of each state. For any edge $e \in \mathcal{E}$ and state $s \in \mathcal{S}$, the state-dependent edge cost function $c^s_e(\cdot)$ is a positive, increasing, and differentiable function of the load through the edge $e$. The state can impact the edge costs in various ways.

The network serves a set of non-atomic travelers with a fixed total demand $D$. We assume that each traveler is subscribed exclusively to one of the TIS in the set $I = \{1, \cdots, I\}$. We refer to the set of travelers subscribed to the TIS $i \in \mathcal{I}$ as population $i$. 

7
All travelers within a population receive an identical signal from their TIS. Let \( \lambda^i \) denote the ratio of population \( i \)'s size and the total demand \( D \). We also consider degenerate situations when the sizes of one or more populations approach 0. Thus, a vector of population sizes \( \lambda = (\lambda^1, \ldots, \lambda^I) \) satisfies \( \sum_{i \in I} \lambda^i = 1 \) and \( \lambda^i \geq 0 \) for any \( i \in I \). The size vector \( \lambda \) is considered as given in our analysis of equilibrium structure and costs (Sections 4 and 5). In Section 6 we consider a more general situation where \( \lambda \) results from the travelers' TIS subscription choices.

Each TIS \( i \in I \) sends a noisy signal \( t^i \) of the state to population \( i \). The signal received by each population determines its type (private information). We assume that the type space of population \( i \) is a finite set, denoted as \( \mathcal{T}^i \). Note that the type spaces \( \mathcal{T}^i \) and the state space \( \mathcal{S} \) need not be of the same size. Let \( t \triangleq (t^1, t^2, \ldots, t^I) \) denote a type profile, i.e. vector of signals received by the traveler populations; thus, \( t \in \mathcal{T} \triangleq \prod_{i \in I} \mathcal{T}^i \).

The joint probability distribution of the state \( s \) and the vector of signals \( t \) is denoted \( \pi \in \Delta(\mathcal{S} \times \mathcal{T}) \), and it is the common prior of the game. The marginal distribution of \( \pi \) on states is consistent with the common prior, i.e. \( \sum_{t \in \mathcal{T}} \pi(s, t) = \theta(s) \) for all \( s \in \mathcal{S} \).

The conditional probability of type profiles \( t \) on the state \( s \) is given by \( p(t|s) = \frac{\pi(s, t)}{\theta(s)} \), i.e. the joint distribution of signals received by the populations when the network state is \( s \).

In our modeling environment, the signals of different TIS can be correlated, conditional on the state. Each population \( i \) generates a belief about the state \( s \) and the other populations' types \( t^{-i} \) based on the signal received from the information system \( i \in I \).

We denote the population \( i \)'s belief as \( \beta^i(s, t^{-i}|t^i) \). The routing strategy of each population \( i \in I \) is a function of its type, denoted as \( q^i(t^i) = (q^i_r(t^i))_{r \in \mathcal{R}} \). One way to describe the generation of routing strategies is that each TIS \( i \in I \) sends a noisy signal \( t^i \) of the state to its subscribed population, and the individual route choices of non-atomic travelers results in an aggregate routing strategy \( q^i(t^i) \).

An alternative viewpoint is that \( q^i(t^i) \) is a direct result of strategy route recommendations sent by each TIS to its subscribed population. That is, each TIS \( i \in I \) routes travelers in population \( i \) according to the function \( q^i(t^i) \). For our purpose, these two viewpoints are equivalent in that given any population \( i \in I \), and any type \( t^i \in \mathcal{T}^i \), the demand of travelers on route \( r \in \mathcal{R} \) is \( q^i_r(t^i) \).

We say that a routing strategy profile \( q \triangleq (q^i)_{i \in I} \) is feasible if it satisfies the following constraints:

\[
\sum_{r \in \mathcal{R}} q^i_r(t^i) = \lambda^i D, \quad \forall t^i \in \mathcal{T}^i, \quad \forall i \in I, \quad (1a)
\]

\[
q^i_r(t^i) \geq 0, \quad \forall r \in \mathcal{R}, \quad \forall t^i \in \mathcal{T}^i, \quad \forall i \in I. \quad (1b)
\]

For a given size vector \( \lambda \), let \( \mathcal{Q}^i(\lambda) \) denote the set of all feasible strategies of population \( i \). From (1a)-(1b), we know that the set of feasible strategy profiles \( \mathcal{Q}(\lambda) \triangleq \prod_{i \in I} \mathcal{Q}^i(\lambda) \) is a convex polytope.
3.2 Bayesian Routing Game

The Bayesian routing game for a fixed size vector $\lambda$ can be defined as

$$\Gamma(\lambda) \triangleq (\mathcal{I}, \mathcal{S}, \mathcal{T}, \mathcal{Q}(\lambda), \mathcal{C}, \beta),$$

where

- $\mathcal{I}$: Set of populations, $\mathcal{I} = \{1, 2, \ldots, I\}$
- $\mathcal{S}$: Set of states with prior distribution $\theta \in \Delta(\mathcal{S})$
- $\mathcal{T} = \prod_{i \in \mathcal{I}} \mathcal{T}^i$: Set of population type profiles with element $t = (t^i)_{i \in \mathcal{I}} \in \mathcal{T}$
- $\mathcal{Q}(\lambda) = \prod_{i \in \mathcal{I}} \mathcal{Q}^i(\lambda)$: Set of feasible strategy profiles for a given size vector $\lambda$, with element $q = (q^i)_{i \in \mathcal{I}} \in \mathcal{Q}(\lambda)$
- $\mathcal{C} = \{c^i_e (\cdot)\}_{e \in \mathcal{E}, s \in \mathcal{S}}$: Set of state-dependent edge cost functions
- $\beta = (\beta^i)_{i \in \mathcal{I}}$: $\beta^i \in \Delta(\mathcal{S} \times \mathcal{T}^{-i})$ is population $i$’s belief on state $s$ and other populations’ types $t^{-i}$

All parameters including the common prior $\pi$ are common knowledge, except that populations privately receive signals about the network state from their respective TIS. The game is played as shown in Fig. 2.

![Figure 2: Timing of the game $\Gamma(\lambda)$.

For any $i \in \mathcal{I}$ and $t^i \in \mathcal{T}^i$, the interim belief of population $i$ is derived from the common prior:

$$\beta^i(s, t^{-i} | t^i) = \frac{\pi(s, t^i, t^{-i})}{\Pr(t^i)}, \quad \forall s \in \mathcal{S}, \quad \forall t^{-i} \in \mathcal{T}^{-i},$$

where $\Pr(t^i) = \sum_{s \in \mathcal{S}} \sum_{t^{-i} \in \mathcal{T}^{-i}} \pi(s, t^i, t^{-i})$. For a strategy profile $q \in \mathcal{Q}(\lambda)$, the induced route flow is denoted $f \triangleq (f_r(t))_{r \in \mathcal{R}, t \in \mathcal{T}}$, where $f_r(t)$ is the aggregate flow assigned to the route $r \in \mathcal{R}$ by populations with type profile $t$, i.e.

$$f_r(t) = \sum_{i \in \mathcal{I}} q^i_r(t^i), \quad \forall r \in \mathcal{R}, \quad \forall t \in \mathcal{T}.$$  

Note that the dependence of $f$ on $q$ is implicit and is dropped for notational convenience.

Again, for the strategy profile $q \in \mathcal{Q}(\lambda)$, we denote the induced edge load as $w \triangleq (w_e(t))_{e \in \mathcal{E}, t \in \mathcal{T}}$, where $w_e(t)$ is the aggregate load on the edge $e$ assigned by populations with type profile $t$:

$$w_e(t) = \sum_{r \in \mathcal{R}} \sum_{i \in \mathcal{I}} q^i_r(t^i) \equiv \sum_{r \in \mathcal{R}} f_r(t), \quad \forall e \in \mathcal{E}, \quad \forall t \in \mathcal{T}.$$
The corresponding cost of edge $e \in E$ in state $s \in S$ is $c^e_d(w_e(t))$. Then, the cost of route $r \in R$ in state $s \in S$ can be obtained as: $c^e_r(q(t)) = \sum_{e \in r} c^e_d(w_e(t))$. Finally, the expected cost of route $r$ for population $i \in I$ can be expressed as follows:

$$
\mathbb{E}[c_r(q)[t^i]] = \sum_{s \in S} \sum_{t^{-i} \in T^{-i}} \sum_{e \in r} \beta^i(s, t^{-i}|t^i)c^e_d(w_e(t^i, t^{-i})), \ \forall r \in R, \ \forall t^i \in T^i, \ \forall i \in I,
$$

(5)

The equilibrium concept for our game $\Gamma(\lambda)$ is Bayesian Wardrop equilibrium (BWE). A strategy profile $q^* \in Q(\lambda)$ is a BWE if for any $i \in I$ and any $t^i \in T^i$:

$$
\forall r \in R, \quad q^i_r(t^i) > 0 \implies \mathbb{E}[c_r(q^*)|t^i] \leq \mathbb{E}[c_r(q^*)|t^i], \ \forall r' \in R.
$$

(6)

That is, in a BWE, travelers in population $i$ with type $t^i$ only take routes that have the smallest expected cost based on their interim belief $\beta^i(s, t^{-i}|t^i)$.

We define the equilibrium population cost, denoted $C^{i*}(\lambda)$, as the expected cost incurred by a traveler of a given population across all types and network states in equilibrium: $C^{i*}(\lambda) \triangleq \frac{1}{\lambda^i} \sum_{t^i \in T^i} \Pr(t^i) \sum_{r \in R} \mathbb{E}[c_r(q^*)|t^i]q^i_r(t^i)$. In fact, from (6), we can write:

$$
C^{i*}(\lambda) \equiv \frac{1}{\lambda^i} \sum_{t^i \in T^i} \Pr(t^i) \left( \sum_{r \in R} q^i_r(t^i) \right) \min_{r \in R} \mathbb{E}[c_r(q^*)|t^i] \equiv \sum_{t^i \in T^i} \Pr(t^i) \min_{r \in R} \mathbb{E}[c_r(q^*)|t^i].
$$

(7)

Note that $\lambda^i = 0$ is a degenerate case for population $i$ as its size approaches 0. In this case, the cost $C^{i*}(\lambda)$ can be viewed as the expected cost faced by an individual (non-atomic) traveler who subscribes to the TIS $i$.

4 Equilibrium Characterization

In this section, we show that the game $\Gamma(\lambda)$ is a weighted potential game. This property enables us to express the sets of equilibrium strategy profiles and route flows as optimal solution sets of certain convex optimization problems.

4.1 Equilibrium Strategy Profiles

Following [Sandholm, 2001], the game $\Gamma(\lambda)$ is a weighted potential game if there exists a continuously differentiable function $\Phi : Q(\lambda) \to \mathbb{R}$ and a set of positive, type-specific weights $\{\gamma(t^i)\}_{t^i \in T^i, i \in I}$ such that:

$$
\frac{\partial \Phi(q(t))}{\partial q^i_r(t^i)} = \gamma(t^i)\mathbb{E}[c_r(q)|t^i], \ \forall r \in R, \ \forall t^i \in T^i, \ \forall i \in I.
$$

(8)

We show that our game $\Gamma(\lambda)$ is a weighted potential game.
Lemma 1. Game $\Gamma(\lambda)$ is a weighted potential game with the potential function $\Phi$ as follows:

$$\Phi(q) \triangleq \sum_{s \in S} \sum_{e \in E} \sum_{t \in T} \pi(s, t) \int_0^{\sum_{r \in R} \sum_{i \in I} q_i(t')} c^s_e(z) dz,$$

and the positive type-specific weight is $\gamma(t^i) = \Pr(t^i)$ for any $t^i \in T^i$ and any $i \in I$.

Using (3) and (4), $\Phi$ can be equivalently expressed as a function of the route flow $f$ or the edge load $w$ induced by a strategy profile $q \in Q(\lambda)$:

$$\hat{\Phi}(f) \triangleq \sum_{s \in S} \sum_{e \in E} \sum_{t \in T} \pi(s, t) \int_0^{\sum_{r \in R} f_r(t)} c^s_e(z) dz$$

$$\hat{\Phi}(w) \triangleq \sum_{s \in S} \sum_{e \in E} \sum_{t \in T} \pi(s, t) \int_0^{w_e(t)} c^s_e(z) dz.$$  \(10\)

Thus, for any feasible strategy profile $q \in Q(\lambda)$, we can write $\Phi(q) \equiv \hat{\Phi}(f) \equiv \hat{\Phi}(w)$, where $f$ and $w$ are the route flow and edge loads induced by the strategy profile $q$. In addition, $\hat{\Phi}(w)$ satisfies the following property:

**Lemma 2.** The function $\hat{\Phi}(w)$ is twice continuously differentiable and strictly convex in $w$.

Our first result provides a characterization of the set of equilibrium strategy profiles:

**Theorem 1.** A strategy profile $q \in Q(\lambda)$ is a BWE if and only if it is an optimal solution of the following convex optimization problem:

$$\min \Phi(q), \quad s.t. \quad q \in Q(\lambda),$$

where $Q(\lambda)$ is the set of feasible strategy profiles. The equilibrium edge load vector $w^*(\lambda)$ is unique.

The existence of BWE follows directly from Theorem 1. For any size vector $\lambda$, we denote the set of BWE for the game $\Gamma(\lambda)$ as $Q^*(\lambda)$. Importantly, since the equilibrium edge load $w^*(\lambda)$ is unique, the equilibrium population cost $C^*(\lambda)$ for each population $i \in I$ in $\Gamma$ must also be unique for any $\lambda$. Thus, the equilibria of $\Gamma(\lambda)$ can be viewed as essentially unique. We denote the optimal value of (OPT-Q), i.e. the value of the weighted potential function $\Phi(q)$ in equilibrium as $\Psi(\lambda)$.

The Lagrangian of (OPT-Q) that we use in proving Theorem 1 is given as follows:

$$L(q, \mu, \nu, \lambda) = \Phi(q) + \sum_{i \in I} \sum_{t^i \in T^i} \mu^{t^i} \left( \lambda^i D - \sum_{r \in R} q^i_r(t^i) \right) - \sum_{r \in R} \sum_{i \in I} \sum_{t^i \in T^i} \nu^i_r q^i_r(t^i),$$

where $\mu = (\mu^{t^i})_{t^i \in T^i, i \in I}$ and $\nu = (\nu^i_r)_{r \in R, t^i \in T^i, i \in I}$ are Lagrange multipliers associated with the constraints (13) and (14), respectively. The next lemma shows that for any BWE $q^* \in Q^*(\lambda)$, the optimal Lagrange multipliers $\mu^*$ and $\nu^*$ in (12) associated with $q^*$ are unique.
Lemma 3. The Lagrange multipliers $\mu^*$ and $\nu^*$ at the optimum of $\text{(OPT-Q)}$ are unique:

$$\mu_{t^i}^* = \min_{r \in R} \Pr(t^i) \mathbb{E}[c_r(q^*)|t^i], \quad \forall t^i \in T^i, \quad \forall i \in I$$  \hspace{1cm} (13a)

$$\nu_{t^i}^{r^i} = \Pr(t^i) \mathbb{E}[c_r(q^*)|t^i] - \mu_{t^i}^*, \quad \forall r \in R, \quad \forall t^i \in T^i, \quad \forall i \in I.$$  \hspace{1cm} (13b)

This result follows from the fact that $\text{(OPT-Q)}$ satisfies the Linear Independence Constraint Qualification (LICQ) condition \cite{Wachsmuth2013}, which ensures the uniqueness of Lagrange multipliers at the optimum of $\text{(OPT-Q)}$; see Lemma A.1 for the statement of LICQ condition.

The value of $\mu_{t^i}^*$ relates the expected route costs for each type $t^i$ in equilibrium with the sensitivity analysis of the Lagrangian with respect to population sizes at the optimum of $\text{(OPT-Q)}$. We will use this result in Section 5 for studying the relative ordering of equilibrium population costs.

4.2 Equilibrium Route Flows

Our main question of interest is how the set of BWE $Q^*(\lambda)$, i.e. optimal solution set of $\text{(OPT-Q)}$, and more importantly, the equilibrium edge load $w^*(\lambda)$, change with the perturbations in the size vector $\lambda$. However, characterizing the effect of $\lambda$ directly from $\text{(OPT-Q)}$ is not so straightforward. Recall that in the simple routing game in Section 2, the effects of perturbations in $\lambda$ on the equilibrium route flow are relatively easier to describe in comparison to the effects on the set of equilibrium strategy profile — the equilibrium route flow remains fixed in a certain range of $\lambda$, whereas the set of equilibrium strategy profiles do not. Thus, our approach involves studying how $\lambda$ affects the set of equilibrium route flows. We show two results in this regard: (i) The set of feasible route flows and the set of feasible strategy profiles that induces a particular route flow can be both expressed as polytopes (Proposition 1); (ii) The set of equilibrium route flows is the optimal solution set of a convex optimization problem (Proposition 2). These results enable us to evaluate how the equilibrium edge load and population costs change with perturbations in $\lambda$.

Let us start by introducing the set of route flows

$$\mathcal{F}(\lambda) \overset{\Delta}{=} \{ f \in \mathbb{R}^{|R| \times |T|} | f \text{ satisfies (14a)-(14d)} \},$$

where the constraints are given by:

$$f_r(t^i, t^{-i}) - f_r(\tilde{t}^i, \tilde{t}^{-i}) = f_r(t^i, \tilde{t}^{-i}) - f_r(\tilde{t}^i, \tilde{t}^{-i}), \quad \forall r \in R, \quad \forall t^i, \tilde{t}^i \in T^i, \text{ and } \forall t^{-i}, \tilde{t}^{-i} \in T^{-i}, \forall i \in I,$$  \hspace{1cm} (14a)

$$\sum_{r \in R} f_r(t) = D, \quad \forall t \in T,$$  \hspace{1cm} (14b)

$$f_r(t) \geq 0, \quad \forall r \in R, \quad \forall t \in T,$$  \hspace{1cm} (14c)

$$D - \sum_{r \in R} \min_{t \in T^i} f_r(t^i, t^{-i}) \leq \lambda^i D, \quad \forall t^{-i} \in T^{-i}, \quad \forall i \in I.$$  \hspace{1cm} (14d)
The constraints (14a)- (14c) do not depend on the size vector \( \lambda \) and can be understood as follows: (14a) captures the fact that the change in the flow through any route resulting from change in the type of population \( i \in I \) does not depend on the particular types of the remaining populations; (14b) ensures that all the demand \( D \) is routed through the network; and (14c) guarantees that the demand assigned to any route is nonnegative.

On the other hand, the constraints in (14d) depend on the size vector \( \lambda \), wherein the size of each population \( i \in I \), \( \lambda^i \), appears linearly in the constraint corresponding to that population. To further interpret (14d), we define the “impact of information” for any given population as the maximum extent to which the signal received from its TIS can influence the routing behavior of travelers within the population. Specifically, for any strategy profile \( q \in Q(\lambda) \) and population \( i \in I \), we define the impact of information on population \( i \in I \) as follows:

\[
J^i(q) \triangleq \lambda^i D - \sum_{r \in R} \min_{t^i \in T^i} q^i_r(t^i) .
\] (15)

Using (1a), we can re-write (13) as:

\[
\hat{J}^i(f) \equiv J^i(q) \equiv \sum_{r \in R} \max_{t^i \in T^i} (f_r(\hat{t}^i, \hat{t}^{-i}) - f_r(t^i, \hat{t}^{-i})) \overset{(14)}{=} D - \sum_{r \in R} \min_{t^i \in T^i} f_r(t^i, \hat{t}^{-i}) ,
\] (16)

where \( (\hat{t}^i, \hat{t}^{-i}) \) is any type profile in \( T \). Now the constraints (14d) can be equivalently stated as:

\[
\hat{J}^i(f) \leq \lambda^i D, \quad \forall i \in I.
\] (IIC)

These constraints ensure that the impact of signals on any population’s strategy is bounded by its size. We will refer to them as information impact constraints (IIC). We use (IIC) and (14d) interchangeably, and refer the constraint in (IIC) corresponding to population \( i \in I \) as (IIC\(_i\)). Also, it is easy to see that for each \( i \in I \), (IIC\(_i\)) can be written as a set of affine inequalities:

\[
D - \sum_{r \in R} f_r(t^i_r, \hat{t}^{-i}) \leq \lambda^i D, \quad \forall t^i_1, \ldots, \forall t^i_{|R|} \in T^i.
\] (18)

Thus, \( F(\lambda) \) is a convex polytope. The following proposition relates the set of feasible strategy profiles and the induced route flows.

**Proposition 1.** The set of feasible route flows is the convex polytope \( F(\lambda) \). Furthermore, for a given route flow \( f \in F(\lambda) \), any feasible strategy profile \( q \in Q(\lambda) \) that induces \( f \) can be expressed as:

\[
q^i_r(t^i) = f_r(t^i, \hat{t}^{-i}) - f_r(\hat{t}^i, \hat{t}^{-i}) + \chi^i_r, \quad \forall r \in R, \quad \forall t^i \in T^i, \quad \forall i \in I,
\] (19)
where $\hat{t} = (\hat{t}_i)_{i \in I}$ is any type profile in $T$, and $\chi = (\chi_r^i)_{r \in R, i \in I}$ satisfies the following constraints:

\begin{align}
\sum_{r \in R} \chi_r^i &= \lambda_i D, \quad \forall i \in I, \quad (20a) \\
\sum_{i \in I} \chi_r^i &= f_r(\hat{t}), \quad \forall r \in R, \quad (20b) \\
\chi_r^i &\geq \max_{t^i \in T^i} (f_r(\hat{t}, \hat{t}^i) - f_r(t^i, \hat{t}^i)), \quad \forall r \in R, \quad \forall i \in I. \quad (20c)
\end{align}

The proof of this proposition is comprised of three steps: In Step I, we show that any route flow that is induced by a feasible strategy profile must satisfy constraints (14a)-(14d). In Step II, we show the converse: for any given $f$, any feasible strategy profile $q$ that induces it must be given by (19), where $\chi$ is a vector satisfying (20). In Step III, we prove that if $f$ satisfies (14a)-(14d), then we can indeed construct a vector $\chi$ that satisfies (20), and the corresponding $q$ in (19) is a feasible strategy profile that induces $f$. By combining Steps II and III, we can conclude that any $f$ satisfying (14a)-(14d) can be induced by at least one feasible strategy profile, and thus is a feasible route flow.

The next proposition provides a characterization of the set of equilibrium route flows, and is analogous to Theorem which characterizes the set of equilibrium strategy profiles.

**Proposition 2.** A feasible route flow $f \in F(\lambda)$ is an equilibrium route flow if and only if $f$ is an optimal solution of the following convex optimization problem:

$$\min \, \hat{\Phi}(f), \quad \text{s.t.} \quad f \in F(\lambda), \quad \text{(OPT-F)}$$

where $F(\lambda)$ is the set of feasible route flow vectors, which satisfy constraints (14a)-(14d).

We denote the set of equilibrium route flows $f^*$ in the game $\Gamma(\lambda)$ as $F^*(\lambda)$. From Theorem equations (11) and (9), we know that for any size vector $\lambda$, and any $q^* \in Q^*(\lambda)$, any $f^* \in F^*(\lambda)$,

$$\Psi(\lambda) = \Phi(q^*) = \hat{\Phi}(f^*) = \hat{\Phi}(w^*(\lambda)). \quad (21)$$

Propositions 1 and 2 form the basis of our analysis of how the perturbations of size vector affects the equilibrium structure and population costs.

## 5 Pairwise Comparison of Populations

In this section, we first analyze the effects of perturbations in the relative sizes of any two populations on the equilibrium structure. Next, we study how the cost difference between any two populations depends on the population sizes.

### 5.1 Equilibrium Regimes

To study the effects of perturbations in the relative sizes of any two populations, we employ the notion of directional perturbation of size vector $\lambda$. In particular, for any two populations $i$ and $j$, we consider the $|I|$-dimensional direction vector $z_{ij} \triangleq (\ldots 0 \ldots, 1, \ldots 0 \ldots, -1,...)$,
...0...) with 1 in the $i$-th position and $-1$ in the $j$-th position. When $\lambda$ is perturbed in the direction of $z^{ij}$, the size of population $i$ (resp. population $j$) increases (resp. decreases), and the sizes of the remaining populations do not change.

For any size vector $\lambda$ and any two populations $i$ and $j$, let the vector of the remaining populations’ sizes be denoted $\lambda^{-ij} \triangleq (\lambda_k)_{k \in \mathcal{I} \setminus \{i, j\}}$. The total size of the remaining populations is $|\lambda^{-ij}| \triangleq \sum_{k \in \mathcal{I} \setminus \{i, j\}} \lambda^k$. For pairwise comparison, we only consider the case when the sizes of both populations are strictly positive so that $|\lambda^{-ij}| < 1$, and the range of the perturbations in the population $i$’s size is $(0, 1 - |\lambda^{-ij}|)$. We denote the set of admissible $\lambda^{-ij}$ as $\Lambda^{-ij}$.

Now consider an optimization problem that is similar to (OPT-$\mathcal{F}$), except that the two constraints in the (IIC) set corresponding to the populations $i$ and $j$ are replaced by a single constraint:

$$\min \; \hat{\Phi}(f), \quad s.t. \; (14a), (14b), (14c), (IIC) \setminus \{i, j\}, (IIC)_{I},$$

where the constraints $(IIC) \setminus \{i, j\}$ indicate that all but $(IIC)$ and $(IIC)_{I}$ from the original set $(IIC)$ are included, and the constraint $(IIC)_{I}$ is defined as follows:

$$\hat{J}^i(f) + \hat{J}^j(f) \leq (1 - |\lambda^{-ij}|)D.$$  \hspace{1cm} (IIC)$_i$

The constraint $(IIC)_{I}$ ensures that the total impact of information on population $i$ and $j$ does not exceed their total demand. We denote the set of optimal solutions for (OPT-$\mathcal{F}^j$) as $\mathcal{F}^{ij,\dagger}$. Analogously to Theorem 1, we can show that any $f^{ij,\dagger} \in \mathcal{F}^{ij,\dagger}$ induces a unique edge load $w^{ij,\dagger}$, which can be obtained by (1); see Lemma B.1. Then, the optimal solution set of (OPT-$\mathcal{F}^j$) can be written as the following polytope:

$$\mathcal{F}^{ij,\dagger} = \left\{ f \left| \begin{array}{c} f \text{satisfies (14a), (14b), (14c), (IIC) \setminus \{i, j\}, and (IIC)_{I},} \\ \sum_{r \in e} f_r(t) = w^{ij,\dagger}_{e}(t), \quad \forall e \in \mathcal{E}, \quad \forall t \in \mathcal{T} \end{array} \right\} \right\}. \hspace{1cm} (22)$$

Note that both $\mathcal{F}^{ij,\dagger}$ and $w^{ij,\dagger}$ depend on $\lambda^{-ij}$ but do not depend on $\lambda^i$ or $\lambda^j$.

Before proceeding further, we need to define two thresholds for the size of one of the two perturbed populations (say, population $i$):

$$\hat{\lambda}^i \triangleq \frac{1}{D} \min_{f^{ij,\dagger} \in \mathcal{F}^{ij,\dagger}} \left\{ \hat{J}^i(f^{ij,\dagger}) \right\}, \quad \hat{\lambda}^i \triangleq \frac{1}{D} \max_{f^{ij,\dagger} \in \mathcal{F}^{ij,\dagger}} \left\{ (1 - |\lambda^{-ij}|)D - \hat{J}^j(f^{ij,\dagger}) \right\},$$

where $\hat{J}^i(f^{ij,\dagger})$ and $\hat{J}^j(f^{ij,\dagger})$ are the impact of information metrics for the population $i$ and $j$, respectively. We can check that $\hat{\lambda}^i$ and $\hat{\lambda}^i$ are admissible thresholds:

**Lemma 4.** $0 \leq \hat{\lambda}^i \leq \hat{\lambda}^i \leq 1 - |\lambda^{-ij}|$.

Additionally, (23) can be expressed as linear programming problems, see (B.7)-(B.8). These two thresholds play a crucial role in our subsequent analysis.

We are now ready to introduce the equilibrium regimes that are induced by the relative change in the sizes of populations $i$ and $j$ with fixed sizes of other populations $\lambda^{-ij} \in \Lambda^{-ij}$. 



15
These regimes are defined by the following sets:

\[
\begin{align*}
\Lambda_1^{ij} & \triangleq \{(\lambda^i, \lambda^j, \lambda^{-ij}) \mid \lambda^i \in (0, \Delta^i)\}, \\
\Lambda_2^{ij} & \triangleq \{(\lambda^i, \lambda^j, \lambda^{-ij}) \mid \lambda^i \in [\Delta^i, \bar{\lambda}^i] \setminus \{0, 1 - |\lambda^{-ij}|\}\}, \\
\Lambda_3^{ij} & \triangleq \{(\lambda^i, \lambda^j, \lambda^{-ij}) \mid \lambda^i \in (\bar{\lambda}^i, 1 - |\lambda^{-ij}|)\}.
\end{align*}
\]

We say that the population \(i\) (resp. population \(j\)) is a “minor population” in regime \(\Lambda_1^{ij}\) (resp. regime \(\Lambda_2^{ij}\)) because \(\lambda^i < \Delta^i\) (resp. \(\lambda^j < 1 - |\lambda^{-ij}| - \bar{\lambda}^i\)). Moreover, neither population is minor in regime \(\Lambda_2^{ij}\). Note that degenerate situations are possible. In particular, if either one or both of the thresholds \(\Delta^i\) and \(\bar{\lambda}^i\) take values in the set \(\{0, 1 - |\lambda^{-ij}|\}\), then from the regime definitions (24a)-(24c), the number of regimes is reduced to two (for example, in the simple routing game in Section 2) or even one regime (see Example B.2). The following theorem describes the properties of equilibrium route flows in the regimes under the directional perturbations in the size vector \(\lambda\).

**Theorem 2.** For any two populations \(i, j \in \mathcal{I}\), and any given \(\lambda^{-ij} \in \Lambda^{-ij}\), the set of equilibrium route flows \(\mathcal{F}^*(\lambda)\) when \(\lambda\) is in regime \(\Lambda_1^{ij}\) or regime \(\Lambda_3^{ij}\) can be expressed as follows:

\[
\mathcal{F}^*(\lambda) = \left\{ \arg \min \Phi(f) \mid \begin{array}{c}
\text{s.t. } \text{(14a), (14b), (14c), (IIC) and } (IIC)\setminus\{j\} \quad \text{if } \lambda \in \Lambda_1^{ij} \\
\text{s.t. } \text{(14a), (14b), (14c), (IIC) and } (IIC)\setminus\{i\} \quad \text{if } \lambda \in \Lambda_3^{ij}
\end{array} \right\}
\]

In regime \(\Lambda_1^{ij}\) (resp. regime \(\Lambda_3^{ij}\)), the constraint (IIC) (resp. (IIC)) is tight in equilibrium. Additionally, in regime \(\Lambda_2^{ij}\), we have \(\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij,\dagger}\).

Essentially, this result is based on how the impact of information on each perturbed population compares with its size; i.e. whether or not the constraint (IIC) (resp. (IIC)) for the population \(i\) (resp. population \(j\)) is tight in equilibrium. In the first side regime \(\Lambda_1^{ij}\), the constraint (IIC) is tight at optimum of (OPT-\(\mathcal{I}\)). This implies that the impact of information extends to the entire demand of the minor population \(i\). In fact, the threshold \(\Delta^i\) is the largest size of population \(i\) for which the impact of information on itself is fully attained. We can argue similarly for the other side regime \(\Lambda_3^{ij}\), where population \(j\) is the minor population; i.e. (IIC\(1\)) is tight at optimum of (OPT-\(\mathcal{I}\)) and \((1 - |\lambda^{-ij}| - \bar{\lambda}^i)\) is the largest size of population \(j\) such that the impact of information on itself is fully attained.

In contrast to the two side regimes, in the middle regime \(\Lambda_2^{ij}\), the sizes of both populations \(i\) and \(j\) are above the threshold sizes \(\Delta^i\) and \((1 - |\lambda^{-ij}| - \bar{\lambda}^i)\), respectively. We can replace the constraints (IIC) and (IIC\(1\)) in the optimization problem (OPT-\(\mathcal{I}\)) by (IIC\(2\)) without changing its optimal value, i.e. the optimal value of (OPT-\(\mathcal{I}\)) is equal to \(\Psi(\lambda)\). However, since the set \(\mathcal{F}^{ij,\dagger}\) (as defined in (22)) contains all route flows that attain the optimal value \(\Psi(\lambda)\) but may not necessarily satisfy the constraints (IIC) and (IIC\(1\)), the equilibrium route flow set \(\mathcal{F}^*(\lambda)\) must be a subset of \(\mathcal{F}^{ij,\dagger}\). In this regime, the impact of information on neither population is fully attained.
A specialized result derived from Theorem 2 and Proposition 1 is that in routing games with two heterogeneously informed populations and a parallel-route network, the equilibrium strategy profile in the regimes $\Lambda_{1}^{ij}$ and $\Lambda_{3}^{ij}$ is unique, see Corollary B.1.

Thanks to Theorem 2, we can analyze the monotonicity of the value of potential function at equilibrium $\Psi(\lambda)$ with respect to perturbations of $\lambda$ in the direction $z^{ij}$.

**Proposition 3.** For any two populations $i, j \in \mathcal{I}$, and any given $\lambda^{-ij} \in \Lambda^{-ij}$, under directional perturbations of $\lambda$ along the direction $z^{ij}$, the function $\Psi(\lambda)$ monotonically decreases in regime $\Lambda_{1}^{ij}$, does not change in $\Lambda_{2}^{ij}$, and monotonically increases in $\Lambda_{3}^{ij}$. Furthermore, the equilibrium edge load vector $w^*(\lambda) = w^{ij;\dagger}$ if and only if $\lambda \in \Lambda_{2}^{ij}$.

Following Theorem 2 in the side regime $\Lambda_{1}^{ij}$ (resp. $\Lambda_{3}^{ij}$), the set of route flows which satisfy the constraints of the optimization problem in (25) increases (resp. decreases) as $\lambda$ is perturbed in the direction $z^{ij}$. Thus, the value of the potential function in equilibrium, $\Psi(\lambda)$, is non-increasing (resp. non-decreasing) in the direction $z^{ij}$. In fact, since the constraint (IIC) (resp. (IIC)) is tight in equilibrium, one can argue that $\Psi(\lambda)$ strictly decreases (resp. increases) in the direction $z^{ij}$. In contrast, in the middle regime $\Lambda_{2}^{ij}$, since $\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij;\dagger}$, we can conclude that $w^*(\lambda) = w^{ij;\dagger}$. Therefore, $\Psi(\lambda) = \hat{\Psi}(w^{ij;\dagger})$, which does not change when $\lambda$ is perturbed in the direction $z^{ij}$.

The necessary and sufficient condition for the invariance of $w^*(\lambda)$ under relative perturbations in the sizes of any two populations in Proposition 3 is a direct consequence of the monotonicity of $\Psi(\lambda)$ and the uniqueness of $w^*(\lambda)$. This result is useful in determining the relative ordering of population costs in equilibrium, as discussed next.

### 5.2 Relative Value of Information

We now study the difference between the equilibrium costs of any two populations under perturbations in their relative sizes. For any two populations $i, j \in \mathcal{I}$ and size vector $\lambda$, we define the relative value of information, denoted $V^{ij*}(\lambda)$, as the expected travel cost saving that a traveler in population $i$ enjoys over a traveler in population $j$, i.e.

$$V^{ij*}(\lambda) \triangleq C^{j*}(\lambda) - C^{i*}(\lambda).$$

Equivalently, $V^{ij*}(\lambda)$ is the expected reduction in the cost faced by an individual traveler when her subscription unilaterally changes from TIS $i$ to TIS $j$, while the TIS subscriptions of all other travelers remain unchanged. We say TIS $i$ is relatively more valuable (resp. less valuable) than TIS $j$ if $V^{ij*}(\lambda) > 0$ (resp. $V^{ij*}(\lambda) < 0$). Similarly, if $V^{ij*}(\lambda) = 0$, TIS $i$ is said to be as valuable as TIS $j$.

It turns out that, for any given size vector $\lambda$, $V^{ij*}(\lambda)$ is closely related to the sensitivity of $\Psi(\lambda)$ (i.e. the value of the potential function in equilibrium) with respect to the perturbation in the relative sizes of populations $i$ and $j$.

**Lemma 5.** The value of the weighted potential function in equilibrium, $\Psi(\lambda)$ as defined in (21), is convex and directionally differentiable in $\lambda$. For any $i, j \in \mathcal{I}$,

$$V^{ij*}(\lambda) = -\frac{1}{D} \nabla_{z^{ij}} \Psi(\lambda),$$

where $\nabla_{z^{ij}} \Psi(\lambda) \triangleq \lim_{\epsilon \to 0^+} \frac{\Psi(\lambda + \epsilon z^{ij}) - \Psi(\lambda)}{\epsilon}$ is the derivative of $\Psi(\lambda)$ in the direction $z^{ij}$. 

17
Proof of Lemma 7. The proof involves applying the results on sensitivity analysis of convex optimization problems, as summarized in Lemmas B.2 and B.3, for a detailed background on these technical results, we refer the reader to [Fiacco 2009], [Fiacco and Kyparisis 1986], and [Rockafellar 1984]. Since in \((\text{OPT}-Q)\), the weighted potential function \(\Phi(q)\) is convex in \(q\), and the constraints \(1a\)–\(1b\) are affine in \(q\) and \(\lambda\), from Lemma B.3 we know that the optimal value of the potential function \(\Psi(\lambda)\) is convex in \(\lambda\).

Next, we can check that \((\text{OPT}-Q)\) satisfies the following conditions: (1) The potential function \(\Phi(q)\) is continuously differentiable in \(q\), and constraints \(1a\)–\(1b\) are unique in equilibrium, \(Q\); (2) The optimal solution set \(Q^*(\lambda)\) is non-empty and bounded (Theorem 1). The Lagrange multipliers at the optimum of \((\text{OPT}-Q)\) are unique, and bounded (Lemma 3). Therefore, from Lemma B.3 we know that \(\Psi(\lambda)\) is differentiable in the direction \(z^{ij}\), and \(\nabla_{z^{ij}} \Psi(\lambda)\) can be expressed as:

\[
\nabla_{z^{ij}} \Psi(\lambda) = \min_{q^* \in Q^*(\lambda)} \max_{(\mu^*, \nu^*)} \nabla_\lambda \mathcal{L}(q^*, \mu^*, \nu^*, \lambda) z^{ij}
\]

where \(M(q^*)\) (resp. \(N(q^*)\)) is the set of optimal Lagrange multipliers \(\mu^*\) (resp. \(\nu^*\)) associated with the equilibrium strategy \(q^*\). From Lemma 3 since both \(\mu^*\) and \(\nu^*\) are unique in equilibrium, \(\nabla_{z^{ij}} \Psi(\lambda)\) can be simplified:

\[
\nabla_{z^{ij}} \Psi(\lambda) = \left( \sum_{t^i \in T^i} \mu^{*t^i} - \sum_{t^j \in T^j} \mu^{*t^j} \right) D
\]

\[
\overset{(1.3a)}{=} \left( \sum_{t^i \in T^i} \min_{r^i \in R} \Pr(t^i) E[c_r(q^*)|t^i] - \sum_{t^j \in T^j} \min_{r^j \in R} \Pr(t^j) E[c_r(q^*)|t^j] \right) D
\]

\[
\overset{7}{=} (C^{i*}(\lambda) - C^{j*}(\lambda)) D = -V^{ij*}(\lambda) \cdot D.
\]

\(\square\)

Our next theorem provides the qualitative structure of relative value of information in the three regimes (24a)–(24c).

**Theorem 3.** For any two populations \(i, j \in \mathcal{I}\), and any \(\lambda^{-ij} \in \Lambda^{-ij}\), the relative value of information \(V^{ij*}(\lambda)\) is positive in regime \(\Lambda^1\), zero in regime \(\Lambda^2\), and negative in regime \(\Lambda^3\).

Furthermore, \(V^{ij*}(\lambda)\) is non-increasing in the direction \(z^{ij}\).

**Proof of Theorem 3.** First, we know from Proposition 3 that in direction \(z^{ij}\), \(\Psi(\lambda)\) decreases in regime \(\Lambda^1\), does not change in regime \(\Lambda^2\) and increases in regime \(\Lambda^3\). Following Lemma 5 we directly obtain that \(V^{ij*}(\lambda) > 0\) in \(\Lambda^1\), \(V^{ij*}(\lambda) = 0\) in \(\Lambda^2\), and \(V^{ij*}(\lambda) < 0\) in \(\Lambda^3\).

Next, from Lemma 5 we know that \(\Psi(\lambda)\) is convex in \(\lambda\). Hence, for any \(i, j \in \mathcal{I}\), the directional derivative \(\nabla_{z^{ij}} \Psi(\lambda)\) is non-decreasing in \(z^{ij}\). From (26), \(V^{ij*}(\lambda)\) is non-increasing in \(z^{ij}\). \(\square\)

Theorem 3 shows that one population has advantage over another population if and only if it is the minor population of the two. For the two side regimes, information impacts
the entire demand of the minor population. As a result, in equilibrium, the travelers in the minor population do not choose the routes with a high expected cost based on the signal they receive from their TIS; however, the travelers in the other population may still choose these routes. On the other hand, in the middle regime, neither population has an advantage over the other one because the information only partially impacts each population’s demand. Consequently, both populations route their demand in a manner such that they face identical cost in equilibrium.

Additionally, the travel cost saving that population $i$ travelers enjoy over the population $j$ is the highest when population $i$ has few travelers. Intuitively, in each side regime, the travelers in the non-minor population face a higher congestion externality relative to the travelers in the minor population, because all travelers within a population are routed according to the same strategy. Naturally, the difference in the equilibrium costs due to the imbalance in congestion externality decreases as the size of the minor population increases, and reduces to zero in the middle regime.

Furthermore, given any two populations $i, j \in I$ and the sizes of all other populations $\lambda^{i,j}$ being fixed, Theorem 3 provides a computational approach to compare the equilibrium costs of populations $i$ and $j$ for the full range of $\lambda^i \in (0, 1 - |\lambda^{i,j}|)$ without explicit computation of BWE or equilibrium route flows for each $\lambda^i$. This approach can be summarized as follows: (i) Solve (OPT-$i^{\lambda}$) to obtain an optimal solution $f^{i^{\lambda},i}$; (ii) Compute $w^{i^{\lambda},i}$ by plugging $f^{i^{\lambda},i}$ into (4); (iii) Obtain $\lambda^i$ and $\bar{\lambda}^i$ by solving (23); and (iv) Find the relative ordering of equilibrium costs of population $i$ and $j$ by checking which of the three possible regimes the size vector $\lambda$ belongs to.

Finally, we can specialize Theorem 3 to analyze situations when a population does not have an access to a TIS, or chooses not to use it at all. Formally, we say that a population $j \in I$ is uninformed if its type is independent with the network state and other populations’ types, i.e. $Pr(t^j|s,t^{-j}) = Pr(t^j)$ for any $t^{-j} \in T^{-j}$, any $t^j \in T^j$, and any $s \in S$. Following (2), the uninformed population $j$’s interim belief can be written as follows:

$$\beta^j(s,t^{-j}|t^j) = \frac{\pi(s,t^j,t^{-j})}{Pr(t^j)} = \frac{Pr(t^j|s,t^{-j}) \cdot Pr(s,t^{-j})}{Pr(t^j)} = Pr(s,t^{-j}) = \sum_{t^j \in T^j} \pi(s,t^{-j},t^j),$$

That is, the interim belief $\beta^j(s,t^{-j}|t^i)$ is identical for any signal $t^i \in T^j$ received by population $j$, and is equal to the marginal probability of $(s,t^{-j})$ calculated from the common prior $\pi$. Therefore, the uninformed population has no further information besides the common knowledge. We show that the equilibrium cost of the uninformed travelers is no less than the cost of any other population.

**Proposition 4.** Consider the game $\Gamma(\lambda)$ in which population $j$ is uninformed. Then, for any size vector $\lambda$, the equilibrium cost of population $j$’s travelers $C^{j*}(\lambda) \geq C^{i*}(\lambda)$, where the population $i$ is any other population (i.e. $i \in I \setminus \{j\}$).

Indeed, if population $j$ is uninformed, we can argue that its equilibrium routing strategy $q^{j*}(t^i)$ must be identical for any $t^i \in T^j$. Consequently, from (15), the impact of
information metric for the population $j$ is $\hat{J}(q^*) = 0$, and perturbing the relative sizes of population $j$ and any other population $i \in \mathcal{I} \setminus \{j\}$ never results in a regime in which population $j$ is the minor population. Applying Theorem 3, we can conclude that the equilibrium cost of population $j$ cannot be less than that of any other population.

We illustrate the results on equilibrium structure and relative value of information in the following two examples:

**Example 1.** We consider a game with two populations on two parallel routes ($r_1$ and $r_2$) with following parameters: $\theta(a) = 0.2$, $D = 10$, $c_1^a(f_1) = f_1 + 15$, $c_1^a(f_1) = 3f_1 + 15$, $c_2(f_2) = 2f_2 + 20$. Types $t^1$ and $t^2$ are independent conditional on the state, i.e. $\Pr(t^1, t^2|s) = \Pr(t^1|s) \cdot \Pr(t^2|s)$. Population 1 has 0.8 chance of getting accurate information of the state, and population 2 has 0.6 chance, i.e. $\Pr(t^1 = s|s) = 0.8$, and $\Pr(t^2 = s|s) = 0.6$. The value of the potential function in equilibrium, equilibrium route flows and population costs are shown in Fig. 3.

In this example, population 1 travelers receive more accurate state information than population 2 travelers. However, population 1 faces a higher cost than population 2 when its size is sufficiently large, i.e., when $\lambda$ is in regime $\Lambda_{3}^{12}$; see Figure 3c. This is due to the fact that in regime $\Lambda_{3}^{12}$, the population 1’s advantage of receiving more accurate information is dominated by the congestion externality it faces due to its relatively large size, in comparison to population 2.

**Example 2.** Let us now consider the game with two populations on two parallel routes ($r_1$ and $r_2$) with the same cost functions, prior distribution $\theta$ and total demand $D$ as that in Example 1. Both populations 1 and 2 have 0.75 chance of getting accurate information about the state, i.e. $\Pr(t^i = s|s) = 0.75$ for any $i \in \mathcal{I}$ and any $s \in \mathcal{S}$. In Fig. 4, we illustrate the equilibrium population costs in two cases: (i) Types $t^1$ and $t^2$ are perfectly correlated, i.e. $t^1 = t^2$; (ii) Types $t^1$ and $t^2$ are independent conditional on the state, i.e. $\Pr(t^1, t^2|s) = \Pr(t^1|s) \cdot \Pr(t^2|s)$.

This example illustrates how the correlation among received signals (or lack thereof) affects the equilibrium structure. Note that case (i) can be viewed as a single-population game. This is because when $t^1$ and $t^2$ are perfectly correlated, there is no information

---

Figure 3: Effects of varying population sizes for Example 1 (a) Weighted potential function in equilibrium, (b) Equilibrium route flows on $r_1$, and (c) Equilibrium population costs.

Figure 4: Effects of varying population sizes for Example 2. (a) Types $t^1$ and $t^2$ are perfectly correlated; (b) Types $t^1$ and $t^2$ are independent conditional on the state.
asymmetry among travelers. Thus, $\lambda^1$ has no impact on the equilibrium outcome, and $\Lambda^1 = 0$, $\bar{\lambda}^1 = 1$ (Fig. 4a). However, case (ii) is not equivalent to a single-population game. Although both populations have identical chance of getting accurate information about the state, there is information heterogeneity among travelers of the two populations, i.e. travelers in one population do not know the signals received by travelers in the other population, and thus the equilibrium outcome changes with the size $\lambda^1$ (Fig. 4b).

We include two additional examples in the e-companion. Example B.1 shows that the regime $\Lambda_{ij}^3$ can be empty even when population 2 is not an uninformed population. Thus, an uninformed population $j$ is sufficient but not necessary for $\bar{\lambda}^i = 1 - |\lambda^{-ij}|$. In Example B.2 we present a situation when only single regime exists in equilibrium.

Our results so far focus on how equilibrium properties and population costs change with the directional perturbation of the size vector $\lambda$. We emphasize that given any $i, j \in \mathcal{I}$, the thresholds $\lambda^i$ and $\bar{\lambda}^i$, as defined in (23), depend on the sizes of the remaining populations $\lambda^{-ij}$, and the populations’ interim beliefs $(\beta^i)_{i \in \mathcal{I}}$ derived from the common prior $\pi$. Importantly, the qualitative structure of the equilibrium regimes resulting from perturbations in the sizes of any two populations is applicable for any size vector $\lambda$ and any common prior. The main property that drives these results is that the equilibrium regimes only depend on whether or not the impact of information on each population is fully attained.

6 General Properties of Equilibrium Outcome

In this section, we first extend our approach of pairwise comparison of populations to study how the equilibrium outcome depends on population sizes in general. Then, we analyze the TIS adoption rates in situations where travelers can choose information subscription.

6.1 Size-Independence of Edge Load Vector

Our analysis in Section 5 showed that if perturbations in the relative sizes of any two populations $i, j \in \mathcal{I}$ induce a middle regime $\Lambda_{ij}^2$, then the equilibrium outcome in this
regime is independent of the sizes of the perturbed populations $i$ and $j$. A natural question to ask is whether this result can be generalized; i.e., can we find a set of size vectors for which the equilibrium edge load does not depend on the size of any population? The answer is affirmative.

We now explicitly characterize the set of size vectors, denoted $\Lambda^\dagger$, for which the edge load is size-independent. Since $(\text{OPT}-\mathcal{F})$ is a convex optimization problem, and $(\text{IIC})$ are the only size-dependent constraints, we can equivalently view $\Lambda^\dagger$ as the set of size vectors for which all the IICs can be dropped from $(\text{OPT}-\mathcal{F})$ without changing its optimal value. Hence, for any $\lambda \in \Lambda^\dagger$, the optimal value of $(\text{OPT}-\mathcal{F})$ is identical to that of the following convex optimization problem:

$$\min \hat{\Phi}(f), \quad \text{s.t.} \quad (14a), (14b) \text{ and } (14c).$$

Let us denote the optimal solution set of (28) as $\mathcal{F}^\dagger$. Analogous to Theorem 1, one can argue that any optimal solution $f^\dagger \in \mathcal{F}^\dagger$ induces a unique edge load $w^\dagger$, obtained from (4). Thus, $\mathcal{F}^\dagger$ can be written as the convex polytope:

$$\mathcal{F}^\dagger = \{ f \mid f \text{ satisfies } (14a), (14b), (14c), \text{ and } \sum_{r \in E} f_r(t) = w^\dagger_r(t), \quad \forall e \in \mathcal{E}, \quad \forall t \in \mathcal{T} \}.$$  

Furthermore, since any route flow in the set $\mathcal{F}^\dagger$ satisfies the constraints $(14a)-(14c)$ – but not necessarily $(\text{IIC})$ constraints – and also attains the optimal value of $(\text{OPT}-\mathcal{F})$, we must have that for any $\lambda \in \Lambda^\dagger$, $\mathcal{F}^\star(\lambda) \subseteq \mathcal{F}^\dagger$. Therefore, for each $\lambda \in \Lambda^\dagger$, there must exist a $f^\dagger \in \mathcal{F}^\dagger$ that is an equilibrium route flow, i.e. at least one $f^\dagger \in \mathcal{F}^\dagger$ satisfies the $(\text{IIC})$ constraints corresponding to $\lambda$:

$$\Lambda^\dagger \Delta \left\{ \lambda \left| \sum_{i \in I} \lambda_i = 1; \quad \lambda_i \geq 0, \quad \forall i \in I; \quad \exists f^\dagger \in \mathcal{F}^\dagger \text{ s.t. } \hat{\mathcal{J}}(f^\dagger) \leq \lambda_i D, \quad \forall i \in I \right. \right\}$$

The following proposition shows the properties of equilibrium edge load and the value of potential function in the set $\Lambda^\dagger$:

**Proposition 5.** The set $\Lambda^\dagger$ is convex, and attains the minimum of $\Psi(\lambda)$, i.e. $\Lambda^\dagger = \arg \min_{\lambda} \Psi(\lambda)$. The equilibrium edge load vector $w^\star(\lambda)$ is size-independent, and is equal to $w^\dagger$ if and only if $\lambda \in \Lambda^\dagger$.

This result shows that some of the properties of $\Psi(\lambda)$ and the change of equilibrium edge load vector under pairwise perturbation (Proposition 3) also hold for the more general case of perturbation in sizes of multiple populations.

### 6.2 Adoption Rates under Choice of TIS

Our analysis so far has focused on the equilibrium properties with fixed population sizes. We now extend our results on the relative value of information (Section 5) and the size independence of the equilibrium edge load vector (Section 6.1) to analyze travelers’ choice of TIS subscription when they can choose to subscribe to any TIS in the set $\mathcal{I}$.
We model travelers’ choice of TIS and the choice of routes as a two-stage game: In the first stage, travelers choose to subscribe to one TIS from the set $\mathcal{I}$. The induced size vector is $\lambda = (\lambda^i)_{i \in \mathcal{I}}$, where $\lambda^i$ is the fraction of travelers who choose TIS $i$. In the second stage, travelers play the Bayesian routing game $\Gamma(\lambda)$. Note that the size vector $\lambda$ here is determined by the travelers’ TIS choices in the first stage, as opposed to being a parameter.

In equilibrium, a traveler who chooses TIS $i$ experiences the expected cost $C^{i*}(\lambda)$. The travelers has no ex-ante incentive to unilaterally change her TIS subscription if and only if $C^{i*}(\lambda)$ is the lowest across all $i \in \mathcal{I}$. Therefore, no traveler has the incentive to change her TIS subscription if and only if:

$$\lambda^i > 0 \quad \Rightarrow \quad C^{i*}(\lambda) = \min_{j \in \mathcal{I}} C^{j*}(\lambda), \quad \forall i \in \mathcal{I}. \quad (31)$$

Such population size vector $\lambda$ can be viewed as the vector of equilibrium adoption rates, one for each TIS. For any size vector that satisfies (31), all travelers experience identical expected costs, and no traveler has the incentive to change her TIS subscriptions.

Our next theorem shows that all size vectors $\lambda \in \Lambda^\dagger$ are equilibrium TIS adoption rates.

**Theorem 4.** The set of equilibrium adoption rates under the choice of TIS is $\Lambda^\dagger$.

**Proof of Theorem 4.** Firstly, we show for any $\lambda \in \Lambda^\dagger$, all travelers have identical costs in equilibrium. Consider any $i, j \in \mathcal{I}$ such that $\lambda^i > 0$ and $\lambda^j > 0$, the directional derivative of $\Psi(\lambda)$ in the direction $z^{ij}$, $\nabla_{z^{ij}} \Psi(\lambda)$, must be 0. Otherwise, $\Psi(\lambda)$ strictly decreases in the direction $z^{ij}$ (resp. $z^{ji}$) if $\nabla_{z^{ij}} \Psi(\lambda) < 0$ (resp. if $\nabla_{z^{ij}} \Psi(\lambda) > 0$), which contradicts the fact that $\Lambda^\dagger = \arg\min_\lambda \Psi(\lambda)$ as in Proposition 3. From Lemma 5, we know that $C^{i*}(\lambda) = C^{j*}(\lambda)$. Therefore, any two populations with positive size have identical costs in equilibrium.

If any $\lambda \in \Lambda^\dagger$ satisfies $\lambda^i > 0$ for all $i \in \mathcal{I}$, then the first step in our proof is sufficient to show that (31) is satisfied. Otherwise, for any $\lambda \in \Lambda^\dagger$, and any degenerate population $i \in \{\mathcal{I} | \lambda^i = 0\}$, we need to show that $C^{i*}(\lambda) \geq C^{j*}(\lambda)$, where $\lambda^j > 0$. Since $\Lambda^\dagger = \arg\min_\lambda \Psi(\lambda)$, we know that $\Psi(\lambda)$ must be non-decreasing in the direction $z^{ij}$. Thus, we obtain: $\nabla_{z^{ij}} \Psi(\lambda) \geq (C^{i*}(\lambda) - C^{j*}(\lambda)) D \geq 0$. Thus, $C^{i*}(\lambda) \geq C^{j*}(\lambda)$. The first and the second steps together show that any $\lambda \in \Lambda^\dagger$ satisfies (31), and hence is a vector of equilibrium adoption rates.

Finally, we show that for any feasible $\lambda \notin \Lambda^\dagger$, (31) is not satisfied. Since $\Lambda^\dagger = \arg\min_\lambda \Psi(\lambda)$, for any $\lambda \notin \Lambda^\dagger$, we can claim that there must exist a direction $z^{ij}$ such that $\Psi(\lambda)$ decreases in the direction $z^{ij}$, $\nabla_{z^{ij}} \Psi(\lambda) < 0$. Otherwise, $\lambda$ is a local minimum of $\Psi(\lambda)$, and since $\Psi(\lambda)$ is convex in $\lambda$ (Lemma 4), $\lambda$ is a global minimum, which contradicts the fact that $\lambda \notin \Lambda^\dagger$. For such a direction $z^{ij}$, there are two possible cases: (1) $\lambda^i > 0$ and $\lambda^j > 0$. In this case, from (26), $C^{i*}(\lambda) \neq C^{j*}(\lambda)$, and thus travelers do not have identical costs in equilibrium. (2) $\lambda^i = 0$ and $\lambda^j > 0$. In this case, from (26), we must have $\nabla_{z^{ij}} \Psi(\lambda) = (C^{i*}(\lambda) - C^{j*}(\lambda)) D < 0$. Therefore, $C^{j*}(\lambda) > C^{i*}(\lambda)$, which implies that travelers in population $j$ has no incentive to change subscription to TIS $i$. To sum up, in either case, $\lambda \notin \Lambda^\dagger$ cannot be a vector of equilibrium adoption rates. \qed
Note that the set $\Lambda^\dagger$ is not a singleton set in general. Recall from Example 1, the set $\Lambda^\dagger$ is the range $\Lambda_1^{\dagger} \cup \Lambda_2^{\dagger}$, in which both TIS are chosen. Therefore, the equilibrium adoption rate of each TIS is not unique. However, since $\Lambda^\dagger$ is a convex set, the equilibrium adoption rate of each TIS $i \in I$ is in a continuous range, denoted $[\lambda_{\text{min}}^i, \lambda_{\text{max}}^i]$, where $\lambda_{\text{min}}^i = \min_{\lambda^i \in \Lambda^\dagger} \lambda^i$ (resp. $\lambda_{\text{max}}^i = \max_{\lambda^i \in \Lambda^\dagger} \lambda^i$) is the minimum (resp. maximum) equilibrium adoption rate. Furthermore, since the set $\Lambda^\dagger$ is determined by the heterogeneous information environment created by all TIS, the equilibrium adoption rate of each TIS is not only determined by the distribution of its own signal, but is also related to the distribution of other TIS signals, and the possible correlations between signals of different TIS.

Finally, Theorem 4 can be used to assess whether or not a set of TIS can induce a heterogeneous information environment. For any $\lambda^i \in \Lambda^\dagger$, the support set of $\lambda^i$, denoted $\bar{I}(\lambda^i) \triangleq \{i \in I | \lambda^i > 0\}$, represents the set of TISs chosen by travelers. In particular, if $|\bar{I}(\lambda^i)| = 1$, then all travelers choose to subscribe to a single TIS even though multiple TIS are available. Thus, the resulting information environment is homogeneous; see Example B.2 where $\lambda^1 = 1$ and $\lambda^2 = 0$ is the only equilibrium adoption rate. However, if $|\bar{I}(\lambda^i)| > 1$, then more than one TIS are chosen, i.e., the heterogeneous information environment is sustained. Moreover, if TIS $i \notin \bar{I}(\lambda^i)$ for any $\lambda^i \in \Lambda^\dagger$, then this TIS is redundant in that it is not chosen in equilibrium even if it is available to travelers.

7 Concluding Remarks

In this article, we study the equilibrium route choices and costs in a heterogeneous information environment, in which each population receives a private signal from their traffic information system (TIS). Each population maintains a belief about the unknown network state and about the signals received by other traveler populations. We focus on analyzing the equilibrium structure under perturbations of population sizes, the relative value of information between any pair of populations, as well as the equilibrium adoption rates when travelers can choose their TIS subscription.

The main ideas behind our analysis approach are: (i) Identification of qualitatively distinct equilibrium regimes based on whether or not the impact of information is fully attained; (ii) Sensitivity analysis of the weighted potential function in equilibrium with respect to the population size vector; and (iii) Characterization of adoption rates under the choice of TIS. Our approach can be easily extended to games where the edge costs are non-decreasing (rather than strictly increasing) in the edge loads. In particular, such a game still admits a weighted potential function, although now the essential uniqueness only applies to the equilibrium edge costs, rather than the loads. The qualitative properties of equilibrium structure, results about the relative ordering of population costs, and adoption rates can be extended as well. However, the characterization of regime thresholds in this case is more complicated from a computational viewpoint due to the non-uniqueness of edge load vector.

One future research question of interest is to analyze how the travelers’ expected cost and TIS adoption rates change when one or more TIS providers make technological changes to their service (for example, improving accuracy levels), or when a new TIS is
introduced. Addressing this problem would involve applying our results to evaluate the value of information for each traveler population as well as the adoption rates under the new information environment, and comparing them with that of the current environment.

Another interesting extension of this work is to study how to design an “optimal” information structure to regulate traffic flows in a network with uncertain state. In the spirit of Bayesian persuasion viewpoint developed by Kamenica and Gentzkow [2011], and Kolotilin et al. [2017], in this extension, an information designer (transportation authority) can send a public or private signal that is correlated with the network state to travelers according to a chosen information structure. The signal determines the (heterogeneous) information environment faced by travelers, and hence influences their routing decisions. The approach presented in this article can be useful in analyzing the impact of information design on travelers’ routing decisions and costs.

Acknowledgement

The authors thank Daron Acemoglu, Tamer Baar, Alexandre Bayen, Sanjeev Goyal, Ali Jadbabaie, Patrick Jaillet, Ramesh Johari, Zaid Khan, Jeffery Liu, Michael Schwarz, Galina Schwartz, Max Shen, David Simchi-Levi, Demos Teneketzis, Pravin Varaiya, Adam Wierman, Muhamet Yildiz for suggestions. We deeply appreciate the feedback of two anonymous reviewers and Associate Editor who reviewed the first version. This work was supported by Google Faculty Research Award, NSF Grant No. 1239054 CPS Frontiers: Foundations Of Resilient CybEr-Physical Systems (FORCES), and NSF CAREER Award CNS 1453126.

References

Daron Acemoglu and Asuman Ozdaglar. Competition and efficiency in congested markets. *Mathematics of Operations Research*, 32(1):1–31, 2007.

Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. Informational Braess paradox: The effect of information on traffic congestion. *Operations Research*, 2018.

Richard Arnott, Andre De Palma, and Robin Lindsey. Does providing information to drivers reduce traffic congestion? *Transportation Research Part A: General*, 25(5):309–318, 1991.

Moshe Ben-Akiva, Andre De Palma, and Kaysi Isam. Dynamic network models and driver information systems. *Transportation Research Part A: General*, 25(5):251–266, 1991.

Moshe Ben-Akiva, Andre de Palma, and Isam Kaysi. The impact of predictive information on guidance efficiency: An analytical approach. In *Advanced Methods in Transportation Analysis*, pages 413–432. Springer, 1996.
David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, pages 265–272, 1953.

José R Correa, Andreas S Schulz, and Nicolás E Stier-Moses. Fast, fair, and efficient flows in networks. *Operations Research*, 55(2):215–225, 2007.

Anthony V Fiacco. Sensitivity and stability in NLP: Continuity and differential stability. *Encyclopedia of Optimization*, pages 3467–3471, 2009.

Anthony V Fiacco and Jerzy Kyparisis. Convexity and concavity properties of the optimal value function in parametric nonlinear programming. *Journal of Optimization Theory and Applications*, 48(1):95–126, 1986.

Anthony V Fiacco et al. *Introduction to sensitivity and stability analysis in nonlinear programming*, volume 90075. Academic press New York, 1983.

Ragavendran Gopalakrishnan, Jason R Marden, and Adam Wierman. Potential games are necessary to ensure pure Nash equilibria in cost sharing games. *Mathematics of Operations Research*, 39(4):1252–1296, 2014.

Ross Haenfler. Social value of public information. *The American Economic Review*, 92(5):1521–1534, 2002.

Jack Hirshleifer. The private and social value of information and the reward to inventive activity. *The American Economic Review*, 61(4):561–574, 1971.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

Zaid Khan and Saurabh Amin. Bottleneck model with heterogeneous information. *Transportation Research Part B: Methodological*, 112:157–190, 2018.

Anton Kolotilin, Tyomfii Mylovanov, Andriy Zapechelnyuk, and Ming Li. Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964, 2017.

Elias Koutsoupias and Christos Papadimitriou. Worst-Case Equilibria. In Christoph Meinel and Sophie Tison, editors, *STACS 99*, volume 1563 of *Lecture Notes in Computer Science*, pages 404–413. Springer Berlin Heidelberg, 1999.

Ehud Lehrer and Dinah Rosenberg. What restrictions do Bayesian games impose on the value of information? *Journal of Mathematical Economics*, 42(3):343–357, 2006.

Hani S Mahmassani and R Jayakrishnan. System performance and user response under real-time information in a congested traffic corridor. *Transportation Research Part A: General*, 25(5):293–307, 1991.

Igal Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13(1):111–124, 1996.
Dov Monderer and Lloyd S Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

Abraham Neyman. The positive value of information. *Games and Economic Behavior*, 3(3):350–355, 1991.

Evdokia Nikolova and Nicolás E Stier-Moses. A mean-risk model for the traffic assignment problem with stochastic travel times. *Operations Research*, 62(2):366–382, 2014.

RT Rockafellar. Directional differentiability of the optimal value function in a nonlinear programming problem. In *Sensitivity, Stability and Parametric Analysis*, pages 213–226. Springer, 1984.

Robert W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, dec 1973. ISSN 0020-7276.

Tim Roughgarden and Éva Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47(2):389–403, 2004.

William H Sandholm. Potential games with continuous player sets. *Journal of Economic theory*, 97(1):81–108, 2001.

Eva Tardos and Tom Wexler. Network formation games and the potential function method. *Algorithmic Game Theory*, pages 487–516, 2007.

Gerd Wachsmuth. On LICQ and the uniqueness of Lagrange multipliers. *Operations Research Letters*, 41(1):78–80, 2013.

Manxi Wu, Jeffrey Liu, and Saurabh Amin. Informational aspects in a class of Bayesian congestion games. In *American Control Conference (ACC), 2017*, pages 3650–3657. IEEE, 2017.
A Supplementary Material for Sec. 4

Proof of Lemma 1. First note that $\Phi(q)$, as defined in (9), is a continuous and differentiable function of the strategy profile $q$. To show that $\Phi(q)$ is a weighted potential function of $\Gamma(\lambda)$, we write the first order derivative of $\Phi(q)$ with respect to $q_e(t')$:

$$\frac{\partial \Phi(q)}{\partial q^i_e(t')} = \sum_{s \in S} \sum_{t - i \in T - i} \pi(s, t', t - i) \sum_{e \in E} c^e_e(w_e(t', t - i))$$

Thus, $\Phi(q)$ satisfies (8) with $\gamma(t') = \Pr(t')$ for any $t' \in T^i$ and any $i \in I$. □

Proof of Lemma 2. Since each $c^e_e(w_e(t))$ is differentiable in $w_e(t)$, we know that $\Phi(w)$ is twice differentiable with respect to $w$. The first order partial derivative of $\Phi(w)$ with respect to $w_e(t)$ can be written as: $\frac{\partial \Phi(w)}{\partial w_e(t)} = \sum_{s \in S} \pi(s, t) c^e_e(w_e(t))$ for any $e \in E$, and any $t \in T$. Also, the second order derivative of $\Phi(w)$ can be written as follows:

$$\frac{\partial^2 \Phi(w)}{\partial w_e(t) \partial w_{e'}(t')} = \begin{cases} \sum_{s \in S} \pi(s, t) \frac{\partial^2 c^e_e(w_e(t))}{\partial w_e(t)} & \text{if } e = e' \text{ and } t = t', \\
0 & \text{otherwise,} \end{cases} \forall e, e' \in E, \forall t, t' \in T.$$

Since for any $e \in E$ and any $s \in S$, $c^e_e(w_e)$ is increasing in $w_e$, $\sum_{s \in S} \pi(s, t) \frac{\partial^2 c^e_e(w_e(t))}{\partial w_e(t)} > 0$.

Thus, the Hessian matrix of $\Phi(w)$ has positive elements on the diagonal and 0 in all other entries, i.e. it is positive definite. Therefore, $\Phi(w)$ is strictly convex in $w$. □

Proof of Theorem 1. We first show that any minimum of (OPT-Q) is a BWE. The Lagrangian of (OPT-Q) is given by (12), where $\mu = (\mu^i_r)_{r \in T^i, i \in I}$ and $\nu = (\nu^i_r)_{r \in R, t' \in T^i, i \in I}$ are Lagrange multipliers associated with the constraints (11a) and (11b), respectively. For any optimal solution $q$, there must exist $\mu$ and $\nu$ such that $(q, \mu, \nu)$ satisfies the following Karush-Kuhn-Tucker (KKT) conditions:

$$\frac{\partial L}{\partial q^i_e(t')} = \frac{\partial \Phi}{\partial q^i_e(t')} - \mu^i_r - \nu^i_r = 0, \forall r \in R, \forall t' \in T^i, \forall i \in I,$$

$$\nu^i_r q^i_e(t') = 0, \forall r \in R, \forall t' \in T^i, \forall i \in I,$$

$$\nu^i_r \geq 0, \forall r \in R, \forall t' \in T^i, \forall i \in I.$$

Using (A.1) and (KKT.1), we have $\frac{\partial \Phi(q)}{\partial q^i_e(t')} = \Pr(t')\mathbb{E}[c_r(q)|t'] = \mu^i_r + \nu^i_r$ for any $r \in R$ and $t' \in T^i, i \in I$. From (KKT.2), we see that for any $r \in R, t' \in T^i, i \in I$, if $q^i_e(t') > 0$, the corresponding Lagrange multiplier $\nu^i_r = 0$, and $\Pr(t')\mathbb{E}[c_r(q)|t'] = \mu^i_r$. However, if $q^i_e(t') = 0$, then $\Pr(t')\mathbb{E}[c_r(q)|t'] = \mu^i_r + \nu^i_r \geq \mu^i_r$. Thus, for any $r \in R, t' \in T^i, i \in I$:

$$q^i_e(t') > 0 \Rightarrow \Pr(t')\mathbb{E}[c_r(q)|t'] = \mu^i_r \leq \mu^i_r + \nu^i_r = \Pr(t')\mathbb{E}[c_r(q)|t'], \forall r' \in R.$$

From (9), we conclude that an optimal solution of (OPT-Q) is a BWE.
Next, we show that any BWE $q^*$ of the game $\Gamma(\lambda)$ is an optimal solution of (OPT-Q). Consider a pair of Lagrange multipliers $\bar{\mu}$ (resp. $\bar{\nu}$) corresponding to the constraints (4a) (resp. (4b)), where $\bar{\mu}^i = \min_{t \in T} \Pr(t^i)E[c_r(q^*)|t^i]$ and $\bar{\nu}^i_t = \Pr(t^i)E[c_r(q^*)|t^i] - \bar{\mu}^i$. We can easily check that (KKT.1) and (KKT.3) are satisfied by $(q^*, \bar{\mu}, \bar{\nu})$. Thus, (KKT.1) and (KKT.3) are satisfied by $(q^*, \bar{\mu}, \bar{\nu})$. Since LICQ holds, for any equilibrium strategy profile $q^*$ in $Q^*$, the induced edge load $\lambda^*(q^*, \bar{\mu}, \bar{\nu})$ is strictly convex in $w$. Therefore, the equilibrium edge load $w^*(\lambda)$ is unique.

**Lemma A.1.** (Theorem 2 in Wachsmuth [2013]) The Lagrange multipliers $\mu^*$ and $\nu^*$ associated with any $q^* \in Q^*(\lambda)$ at the optimum of (OPT-Q) are unique if and only if the LICQ condition is satisfied in that the gradients of the set of tight constraints in (1a)–(1b) at the optimum are linearly independent.

**Proof of Lemma A.1.** Let the set of constraints that are tight at optimum of (OPT-Q) in (1b) be denoted as $B$. Assume for the sake of contradiction that LICQ does not hold, i.e. the set of equality constraints (1a) and the elements in the set $B$ are linearly dependent. Now, note that the constraint sets (1a) and (1b) are each comprised of linearly independent affine functions. Hence, there must exist a type $\bar{r}$ such that the gradient of the corresponding equality constraint (i.e. $\sum_{r \in R} q^*_r(\bar{r}) = \lambda^*D$) is linearly dependent with the elements in the set $B$, which implies that $q^*_r(\bar{r}) = 0, \forall r \in R$. However, this violates the equality constraint in (1a) as $\sum_{r \in R} q^*_r(\bar{r}) = \lambda^*D \neq 0$, and we arrive at a contradiction.

Since LICQ holds, for any equilibrium strategy profile $q^* \in Q^*(\lambda)$, the corresponding $\mu^*$ and $\nu^*$ must be unique. Following the proof of Theorem 1, we conclude that for any $q^* \in Q^*(\lambda)$, $(q^*, \mu^*, \nu^*)$ satisfies the KKT conditions, and $\mu^t*$ and $\nu^t*$ can be written as (13a) and (13b), respectively.

Finally, noting that the equilibrium edge load is unique (Theorem 1), $\mu^*$ and $\nu^*$ in (13a)–(13b) are thus unique in equilibrium.

**Proof of Proposition 1.** Step 1: We show that any $q \in Q(\lambda)$ induces a route flow $f$ that satisfies (14a)–(14d). From (3), we obtain that for any $t^i, \bar{t}^i \in T^i$, any $t^{-i}, \bar{t}^{-i} \in T^{-i}$,
and any \( i \in I \), \( f \) satisfies (14a):

\[
 f_r(t^i, t^{-i}) - f_r(\tilde{t}^i, t^{-i}) = q^i_r(t^i) + \sum_{j \in I \setminus \{i\}} q^i_j(t^j) - q^i_j(\tilde{t}^j) - \sum_{j \in I \setminus \{i\}} q^i_j(t^j)
\]

\[
 = q^i_r(t^i) + \sum_{j \in I \setminus \{i\}} q^i_j(\tilde{t}^j) - q^i_j(t^j) - \sum_{j \in I \setminus \{i\}} q^i_j(\tilde{t}^j) = f_r(t^i, \tilde{t}^{-i}) - f_r(\tilde{t}^i, \tilde{t}^{-i}).
\]

From (1a) and (1b), we can directly conclude that \( f \) must also satisfy (14b) and (14c). Additionally,

\[
 D - \sum_{r \in R} \min_{t \in T_r} f_r(t^i, t^{-i}) \overset{1a}{=} D - \sum_{r \in R} \left( \sum_{j \in I \setminus \{i\}} q^i_j(t^j) - \min_{t \in T_r} q^i_j(t^j) \right) \overset{1b}{=} D - \sum_{r \in R} \lambda^i D - \sum_{t \in T_r} \min_{r \in R} q^i_r(t^r) \overset{(13)}{=} \lambda^i D, \quad \forall t^{-i} \in T^{-i}, \forall i \in I.
\]

Therefore, \( f \) satisfies (14d). Thus, any feasible route flow must satisfy (14a)-(14d).

Step II: Next, we show that for any route flow \( f \in F(\lambda) \) (i.e. \( f \) that satisfies constraints (14a)-(14d)), the set of feasible strategies that induce \( f \) can be characterized by (19). For any route \( r \in R \), the linear system of equations (3) has \( \prod_{i \in I} |T^i| \) equations in \( \sum_{i \in I} |T^i| \) variables. Note that for any given \( \tilde{t} = (\tilde{t}^i)_{i \in I} \in T \), the following equations are linearly independent:

\[
 \sum_{i \in I} q^i_r(\tilde{t}^i) = f_r(\tilde{t}),
\]

\[
 q^i_r(t^i) + \sum_{j \in I \setminus \{i\}} q^i_j(\tilde{t}^j) = f_r(t^i, \tilde{t}^{-i}), \quad \forall t^i \in T^i \setminus \{\tilde{t}^i\}, \quad \forall i \in I.
\]

We then show that given any \( t \in T \), \( \sum_{i \in I} q^i_r(t^i) = f_r(t) \) is a linear combination of the equations in (A.3). Following (14a), we can write:

\[
 \sum_{i \in I} f_r(t^i, \tilde{t}^{-i}) - (|I| - 1)f_r(\tilde{t}) = f_r(t^1, \tilde{t}^{-1}) + f_r(t^2, \tilde{t}^{-2}) + \sum_{i=3}^{|I|} f_r(t^i, \tilde{t}^{-i}) - (|I| - 1)f_r(\tilde{t}) \overset{1m}{=} f_r(t^1, t^2, \tilde{t}^{-1-2}) + f_r(\tilde{t}) + \sum_{i=3}^{|I|} f_r(t^i, \tilde{t}^{-i}) - (|I| - 1)f_r(\tilde{t})
\]

\[
 = f_r(t^1, t^2, \tilde{t}^{-1-2}) + \sum_{i=3}^{|I|} f_r(t^i, \tilde{t}^{-i}) - (|I| - 2)f_r(\tilde{t}),
\]

where \( \tilde{t}^{-1-2} = (\tilde{t}^3, \ldots, \tilde{t}^{|I|}) \). We apply the same procedure iteratively for another \( |I| - 2 \) times:

\[
 \sum_{i \in I} f_r(t^i, \tilde{t}^{-i}) - (|I| - 1)f_r(\tilde{t}) = f_r(t), \quad \forall t \in T.
\]
Now for any \( r \in \mathcal{R} \) and \( t \in \mathcal{T} \), we can write \( \sum_{i \in I} q^i_r(t') = \sum_{i \in I} \left( q^i_r(t') + \sum_{j \in I \setminus \{i\}} q^j_r(t') \right) - (|I| - 1) \sum_{i \in I} q^i_r(t') \equiv \sum_{i \in I} f_r(t', \hat{t} - i) - (|I| - 1) f_r(t) \equiv f_r(t). \) Thus, for any \( r \in \mathcal{R} \), the linear system (3) is comprised of \( \sum_{i \in I} \mathcal{T}^i \) variables, and any constraint can indeed be expressed as a linear combination of \( \sum_{i \in I} \mathcal{T}^i - |I| + 1 \) independent equations in (A.3). From the rank-nullity theorem, we conclude that the dimension of null space of this linear map is \( |I| - 1 \). Then, for any \( r \in \mathcal{R} \), any \( i \in \mathcal{I} \), setting \( q^i_r(t') = \chi^i_r \), any solution of (3) can be expressed as (19), where \( \chi^i_r \) is an arbitrary type profile. Additionally, \( \sum_{i \in I} \chi^i_r = \sum_{i \in I} q^i_r(t') = f_r(\hat{t}). \) Thus, \( \chi \) satisfies (20b), i.e. for each \( r \in \mathcal{R} \), out of the \( |I| \) variables in \( \{\chi^i_r\}_{i \in \mathcal{I}} \), \( |I| - 1 \) are free, and the remaining one is obtained from (20b). We can conclude that the strategy profile \( q \) as defined in (19) induces the route flow \( f \). It remains to be shown that if \( q \) is a feasible strategy profile, \( \chi \) must satisfy (20a) and (20c) as well. Since \( q \) satisfies (1a), we obtain that \( \chi^i_r = \frac{\mathcal{X}^i_r - \sum_{r \in \mathcal{R}} q^i_r(t')}{|I| - 1} \sum_{r \in \mathcal{R}} \left( f_r(t', \hat{t} - i) - f_r(\hat{t}, \hat{t} - i) + \chi^i_r \right) \). Additionally, from (1b), \( 0 \leq q^i_r(t') = f_r(t', \hat{t} - i) - f_r(\hat{t}, \hat{t} - i) + \chi^i_r \) for any \( r \in \mathcal{R} \) and any \( t \in \mathcal{T} \). Thus, \( \chi^i_r \geq \max_{t \in \mathcal{T}} \{ f_r(\hat{t}, \hat{t} - i) - f_r(t', \hat{t} - i) \} \), i.e. \( \chi \) satisfies (20c).

Step III: Finally, we show that the set of \( \chi \) satisfying (20) is non-empty, i.e., any \( f \in \mathcal{F}(\lambda) \) can be induced by at least one feasible strategy profile \( q \). Consider any \( f \in \mathcal{F}(\lambda) \), we explicitly construct the following \( \chi \), and show that such \( \chi \) satisfies (20):

\[
\chi^i_r = \gamma^i_r \left( \lambda^i_D - \sum_{r \in \mathcal{R}} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i)) \right) + \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i)) \quad \forall r \in \mathcal{R}, \forall i \in \mathcal{I}, \tag{A.5}
\]

where \( \hat{t} \) is any arbitrary type profile, and \( \gamma^i_r = \frac{f_r(t') - \sum_{i \in I} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i))}{\sum_{r \in \mathcal{R}} \[ f_r(t') - \sum_{i \in I} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i))] - \sum_{i \in I} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i))} \times \frac{1}{|I| - 1} \). Note that \( \gamma^i_r \geq 0 \). To see this, since for any \( t \in \mathcal{T} \), \( \sum_{i \in I} f_r(t', \hat{t} - i) - (|I| - 1) f_r(t) \equiv f_r(t) \geq 0 \), we know that \( \min_{t' \in \mathcal{T}} \sum_{i \in I} f_r(t', \hat{t} - i) - (|I| - 1) f_r(t) = \min_{t' \in \mathcal{T}} f_r(t') \geq 0 \). Thus, for any \( r \in \mathcal{R} \), we obtain:

\[
f_r(t') - \sum_{i \in I} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i)) = \min_{t \in \mathcal{T}} \sum_{i \in I} f_r(t', \hat{t} - i) - (|I| - 1) f_r(t) = \min_{t \in \mathcal{T}} f_r(t) \geq 0. \tag{A.6}
\]

Hence, we can conclude that \( \gamma^i_r \geq 0 \). Next, \( \lambda^i_D - \sum_{r \in \mathcal{R}} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i)) \equiv \lambda^i_D - \left( \sum_{r \in \mathcal{R}} \min_{t' \in \mathcal{T}} f_r(t', \hat{t} - i) \right) \geq 0 \). Using the above inequalities, we obtain that \( \chi^i_r \) as considered in (A.5) satisfies (20c).

Second, we check \( \chi^i_r \) satisfies (20a). If \( \sum_{r \in \mathcal{R}} \left[ f_r(t') - \sum_{i \in I} \max_{t' \in \mathcal{T}} (f_r(t') - f_r(t', \hat{t} - i)) \right] > 31 \)
which implies that for any $i \neq 0$, then:

\[
\sum_{r \in \mathcal{R}} \chi^i_r = \sum_{r \in \mathcal{R}} \gamma_r \left( \lambda^D \sum_{i' \in \mathcal{T}^i} \max_{r' \in \mathcal{R}} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) \right) + \sum_{r \in \mathcal{R}} \max_{i' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) = \lambda^D D.
\]

On the other hand, if $\sum_{r \in \mathcal{R}} [f_r(\hat{t}) - \sum_{i' \in \mathcal{T}^i} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i}))] = 0$, we obtain that:

\[
0 = \sum_{r \in \mathcal{R}} [f_r(\hat{t}) - \sum_{i' \in \mathcal{T}^i} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i}))] \overset{(\text{14a})}{=} D - \sum_{i' \in \mathcal{T}^i} \left( \sum_{r \in \mathcal{R}} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) \right) \geq D - \sum_{i' \in \mathcal{T}^i} \lambda^D D = 0,
\]

which implies that for any $i \in \mathcal{I}$, $\sum_{r \in \mathcal{R}} \max_{i' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) = \lambda^D D$. Since in this case, $\gamma_r = 0$, we can conclude that $\sum_{r \in \mathcal{R}} \chi^i_r = \sum_{r \in \mathcal{R}} \max_{i' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) = \lambda^D D$, i.e. $\chi$ satisfies (20a).

Finally, $\chi^i_r$ also satisfies (20b). If $\sum_{r \in \mathcal{R}} [f_r(\hat{t}) - \sum_{i' \in \mathcal{T}^i} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i}))] > 0$, we have:

\[
\sum_{i' \in \mathcal{T}^i} \max_{i' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) = f_r(\hat{t}) - \sum_{i' \in \mathcal{T}^i} \max_{i' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) + \sum_{i' \in \mathcal{T}^i} \max_{i' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) = f_r(\hat{t}).
\]

If $\sum_{r \in \mathcal{R}} [f_r(\hat{t}) - \sum_{i' \in \mathcal{T}^i} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i}))] = 0$, then we have

\[
0 = \sum_{r \in \mathcal{R}} [f_r(\hat{t}) - \sum_{i' \in \mathcal{T}^i} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i}))] \overset{(\text{A.6})}{=} \sum_{r \in \mathcal{T}^i} \min_{t \in \mathcal{T}^i} f_r(t) \geq 0,
\]

which implies that for any $r \in \mathcal{R}$, $\min_{t \in \mathcal{T}^i} f_r(t) = 0$. In this case, $\gamma_r = 0$, and thus $\sum_{i' \in \mathcal{T}^i} \chi^i_r = \sum_{i' \in \mathcal{T}^i} \max_{t' \in \mathcal{T}^i} (f_r(\hat{t}) - f_r(t^i, \hat{t}^{-i})) \overset{(\text{A.6})}{=} f_r(\hat{t}) - \min_{t \in \mathcal{T}^i} f_r(t) = f_r(\hat{t})$, i.e. $\chi$ satisfies (20b).

Therefore, if $f$ satisfies (14a)-(14d), we can conclude that the $\chi$ in (A.5) satisfies (20), i.e. the set of $\chi$ satisfying (20) is non-empty. We already showed in Step II that $q$ as defined in (19) with parameter $\chi$ satisfying (20) is a feasible strategy profile, and $q$ induces
Therefore, if $f$ satisfies (14a)-(14d), there exists a feasible $q$ that induces $f$, i.e. any $f \in \mathcal{F}(\lambda)$ is a feasible route flow.

In summary, we have shown that any feasible route flow satisfies (14) (Step I); For any $f$ that satisfies (14), the set of feasible strategy profiles that induce $f$ can be written in (19)-(20) (Step II); Such set is non-empty, and hence $f$ is feasible (Step III). We can thus conclude that the set of feasible route flows is $\mathcal{F}(\lambda)$, and the set of feasible strategies that induce $f$ is as in (19)-(20). $\square$

**Proof of Proposition 3.** From Proposition 1 we know that the set of feasible route flows is the set $\mathcal{F}^\mathcal{Q}(\lambda)$ characterized by (14a)-(14d). Additionally, the weighted potential function in (9) can be equivalently written as a function of $f$ given by (10). Therefore, the minimum of $(\text{OPT} - \mathcal{F}^\mathcal{Q})$ is equal to that in $(\text{OPT} - \mathcal{Q})$, and the set of optimal solutions is the set of equilibrium route flows. $\square$

### Supplementary Material for Section 5

**Lemma B.1.** The route flows $f^{ij, \dagger} \in \mathcal{F}^{ij, \dagger}$ induce a unique edge load $w^{ij, \dagger}$.

**Proof of Lemma B.1.** Following (4) and (OPT - $\mathcal{F}^{ij, \dagger}$), any edge load $w^{ij, \dagger}$ induced by route flows in $\mathcal{F}^{ij, \dagger}$ (which we defined as optimal solution set of (OPT - $\mathcal{F}^{ij}$)) is an optimal solution of the following optimization problem:

$$\min_w \tilde{\Phi}(w),$$

$$\text{s.t. } w_e(t) = \sum_{r \in e} f_r(t), \quad \forall t \in \mathcal{T}, \quad \forall e \in \mathcal{E},$$

$$f \text{ satisfies (14a), (14b), (14c), (IIC)} \{i, j\}, \text{ (IIC)}^{ij}.$$  

The constraints (14a), (14b), (14c) are linear constraints. Following from (18), constraints (IIC) \{i, j\}, (IIC$^{ij}$) are each equivalent to a set of linear constraints. Additionally, $w$ is a linear function of $f$, thus the feasible set of $w$ in this optimization problem must also be a convex polytope. Since $\tilde{\Phi}(w)$ is a strictly convex function in $w$, the optimal solution $w^{ij, \dagger}$ is unique. $\square$

**Proof of Lemma 4.** First, we show that both thresholds $\lambda_i^\mathcal{Q}$ and $\bar{\lambda}_i^\mathcal{Q}$ belong to the interval $[0, 1 - |\lambda^{-ij}|]$. Since $\lambda_i^\mathcal{Q}$ is attainable on the set $\mathcal{F}^{ij, \dagger}$, there exists $\bar{f}^{ij, \dagger} \in \mathcal{F}^{ij, \dagger}$ such that:

$$\lambda_i^\mathcal{Q} = \frac{1}{D} \tilde{J}(\bar{f}^{ij, \dagger}) \overset{\text{(10)}}{=} \frac{1}{D} \left(D - \sum_{r \in \mathcal{R}} \min_{t \in \mathcal{T}} \bar{f}^{ij, \dagger}_r (t^i, t^{-i}) \right) \geq \frac{1}{D} \left(D - \sum_{r \in \mathcal{R}} \tilde{f}^{ij, \dagger}_r (\tilde{t}^i, t^{-i}) \right) \overset{\text{(14d)}}{=} 0.$$  

Similarly, we can check that $\bar{\lambda}_i^\mathcal{Q} \leq 1 - |\lambda^{-ij}|$.

Additionally, we know for any $\bar{f}^{ij, \dagger} \in \mathcal{F}^{ij, \dagger}$:

$$\bar{\lambda}_i^\mathcal{Q} \overset{\text{(20)}}{=} \frac{1}{D} \left\{ (1 - |\lambda^{-ij}|) D - \tilde{J}(\bar{f}^{ij, \dagger}) \right\} \overset{\text{(IIC)}}{\geq} \frac{1}{D} \tilde{J}(\bar{f}^{ij, \dagger}) \overset{\text{(20)}}{\geq} \lambda_i^\mathcal{Q}.$$  

Therefore, $0 \leq \lambda_i^\mathcal{Q} \leq \bar{\lambda}_i^\mathcal{Q} \leq 1 - |\lambda^{-ij}|$. $\square$
For any two populations $i, j \in \mathcal{I}$, we can compute the threshold $\lambda^i$ as follows:

$$
\begin{align*}
\min & \quad y \\
\text{s.t.} & \quad D - \sum_{r \in \mathcal{R}} f_r(t^i_r, \hat{r}^{-i}) \leq y \cdot D, \quad \forall t^i_1 \in \mathcal{T}^i, \ldots, \forall t^i_{|\mathcal{R}|} \in \mathcal{T}^i, \\
& \quad f^{ij, \dagger} \in \mathcal{F}^{ij, \dagger},
\end{align*}
$$

where $\mathcal{F}^{ij, \dagger}$ is the polytope defined in (22). Therefore, (B.7) is a linear programming. Analogously, the threshold $\bar{\lambda}^i$ is the optimal value of the following linear program:

$$
\begin{align*}
\max & \quad y \\
\text{s.t.} & \quad -|\lambda^{-ij}|D + \sum_{r \in \mathcal{R}} f_r(t^i_r, \hat{r}^{-j}) \geq y \cdot D, \quad \forall t^j_1 \in \mathcal{T}^j, \ldots, \forall t^j_{|\mathcal{R}|} \in \mathcal{T}^j, \\
& \quad f^{ij, \dagger} \in \mathcal{F}^{ij, \dagger}.
\end{align*}
$$

Proof of Theorem 2 [Regime $\Lambda^j_1$]: First, we show by contradiction that the constraint $\text{(IC)}$ is tight for any equilibrium route flow. Assume that for a given $\lambda \in \Lambda^j_1$, there exists an equilibrium route flow $f^*$ such that $\text{(IC)}$ is not tight. From Proposition 2, we know that $f^*$ is an optimal solution of (OPT-$F^j$). Since (OPT-$F$) is a convex optimization problem, $f^*$ is still a minimizer of $\hat{\Phi}(f)$ if we drop the constraint $\text{(IC)}$. Additionally, the constraints $\text{(IC)}$ and $\text{(IC)_j}$ implies that $f^*$ must also satisfy $\text{(IC)_j}$. Thus, such $f^*$ is an optimal solution of the following problem:

$$
\begin{align*}
\min & \quad \hat{\Phi}(f), \\
\text{s.t.} & \quad (14a), (14b), (14c), (\text{IC}_j), \text{and } (\text{IC}) \setminus \{i\}.
\end{align*}
$$

Moreover, the threshold $\bar{\lambda}^i$ defined in (23) is attained by a route flow, say $\tilde{f}^{ij, \dagger}$, in the set $\mathcal{F}^{ij, \dagger}$. Thus, we can write: $1 - |\lambda^{-ij}| - \frac{1}{D} \tilde{j}^i(\tilde{f}^{ij, \dagger}) = \bar{\lambda}^i \geq \lambda^i \geq \lambda^i$. Rearranging, we obtain: $\frac{1}{D} \tilde{j}^i(\tilde{f}^{ij, \dagger}) < 1 - |\lambda^{-ij}| - \lambda^i = \lambda^i$, and so such $\tilde{f}^{ij, \dagger}$ also satisfies $\text{(IC)}$. Since $\tilde{f}^{ij, \dagger}$ is an optimal solution of (22), which minimizes the same objective function as (B.9) but without the constraint $\text{(IC)_j}$, we thus know that $\tilde{f}^{ij, \dagger}$ is also an optimal solution in (B.9). Since the induced edge load is unique, we can conclude that the edge load induced by $f^*$ must be identical to that induced by $\tilde{f}^{ij, \dagger}$, which is $w^{ij, \dagger}$. Then, from (22), we have $f^* \in \mathcal{F}^{ij, \dagger}$. Therefore, from (23), we can write $\lambda^i \leq \frac{1}{D} \tilde{j}^i(f^*)$. Since we assumed that $\text{(IC)}$ is not binding in equilibrium, we obtain: $\frac{1}{D} \tilde{j}^i(f^*) < \lambda^i < \lambda$ $\leq \frac{1}{D} \tilde{j}^i(f^*)$, which is a contradiction. Thus, $\text{(IC)}$ must be tight in equilibrium for any $\lambda$ in regime $\Lambda^j_1$.

Finally, following the tightness of $\text{(IC)}$ at optimum of (OPT-$F$), by rearranging the constraint $\text{(IC)_j}$ in (25), we have $\tilde{j}^i(f^*) \leq (1 - |\lambda^{-ij}|)D - \tilde{j}^i(f^*) = \lambda D$. Thus, $\text{(IC)_j}$ is guaranteed to hold in Regime $\Lambda^j_1$ given the constraint $\text{(IC)_j}$ and the fact that $\text{(IC)}$ is tight at the optimum of (OPT-$F$). Hence, $\text{(IC)}$ can be dropped in (OPT-$F$) without changing the optimal solution set.

[Regime $\Lambda^j_2$]: Analogous to the proof given for regime $\Lambda^j_1$, we can argue that constraint $\text{(IC)}$ is tight in any equilibrium for any $\lambda$ in regime $\Lambda^j_2$. By imposing constraint
\((\text{IIC}_1)\), \((\text{IIC}_3)\) can be dropped from the constraint set in \((\text{OPT}-\mathcal{F})\) without changing the optimal solution set.

[Regime \(\Lambda_2^\downarrow\)]: To study this regime, we need two additional thresholds

\[
\lambda^i \triangleq \frac{1}{D} \max_{f^{ij \downarrow} \in \mathcal{F}^{ij \downarrow}} \left\{ \hat{J}_i(f^{ij \downarrow}) \right\}, \quad \text{and} \quad \bar{\lambda}^i \triangleq \frac{1}{D} \min_{f^{ij \uparrow} \in \mathcal{F}^{ij \uparrow}} \left\{ (1 - |\lambda - \lambda^i|) D - \hat{J}_i(f^{ij \uparrow}) \right\}.
\]

From (23), we can check that \(\lambda^i \leq \underline{\lambda}^i\), and \(\bar{\lambda}^i \leq \bar{\lambda}^i\).

For any \(\lambda^i \in [\underline{\lambda}^i, \bar{\lambda}^i]\), we argue that \(\mathcal{F}^i(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\). Since the set \(\mathcal{F}^{ij \downarrow}\) as defined by (22) is a bounded polytope, and \(\lambda^i\) (resp. \(\underline{\lambda}^i\)) is the minimum (resp. maximum) value of the continuous function \(\hat{J}_i(f^{ij \downarrow})\) on \(\mathcal{F}^{ij \downarrow}\), we know from the mean value theorem that there exists a \(\hat{f}^{ij \downarrow} \in \mathcal{F}^{ij \downarrow}\) satisfying: \(\lambda^i = \frac{1}{D} \hat{J}_i(\hat{f}^{ij \downarrow})\). Such \(\hat{f}^{ij \downarrow}\) also satisfies constraint \((\text{IIC}_1)\). Therefore, \(\hat{f}^{ij \downarrow}\) satisfies all the constraints in (14), and minimizes \(\hat{\Phi}(f)\). So \(\hat{f}^{ij \downarrow}\) is an equilibrium route flow, which implies that \(\mathcal{F}^*(\lambda) \cap \mathcal{F}^{ij \downarrow} \neq \emptyset\). Since the equilibrium edge load vector is unique, and the edge load induced by \(\hat{f}^{ij \downarrow}\) is \(w^{ij \downarrow}\), we must have \(w^*(\lambda) = w^{ij \downarrow}\). Furthermore, from (22), \(\mathcal{F}^{ij \downarrow}\) is a superset of all feasible route flows that can induce \(w^{ij \downarrow}\). Therefore, \(\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\) for any \(\lambda^i \in [\underline{\lambda}^i, \bar{\lambda}^i]\). Similarly, we can argue that for any \(\lambda^i \in [\underline{\lambda}^i, \bar{\lambda}^i]\), \(\mathcal{F}^i(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\).

To prove that \(\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\) for any \(\lambda\) in regime \(\Lambda_2^\downarrow\), we need to argue two cases \(\underline{\lambda}^i \geq \bar{\lambda}^i\) and \(\underline{\lambda}^i < \bar{\lambda}^i\) separately. If \(\underline{\lambda}^i \geq \bar{\lambda}^i\), then \([\underline{\lambda}^i, \bar{\lambda}^i] \subseteq [\underline{\lambda}^i, \bar{\lambda}^i] \cup [\bar{\lambda}^i, \underline{\lambda}^i]\). Therefore, \(\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\) for any \(\lambda^i \in (\underline{\lambda}^i, \bar{\lambda}^i)\), we can check that any \(f^{ij \downarrow} \in \mathcal{F}^{ij \downarrow}\) satisfies the constraint \((\text{IIC}_1)\): \(\frac{1}{D} \hat{J}_i(f^{ij \downarrow}) \leq \lambda^i < \lambda^i\). Additionally, since \(\lambda^i < \bar{\lambda}^i < 1 - |\lambda - \lambda^i| - \frac{1}{D} \hat{J}_i(f^{ij \downarrow})\), we know that \(\frac{1}{D} \hat{J}_i(f^{ij \downarrow}) < 1 - |\lambda - \lambda^i| - \lambda^i = \lambda^i\), i.e. \(f^{ij \downarrow}\) also satisfies the constraint \((\text{IIC}_1)\). Thus, any \(\hat{f}^{ij \downarrow} \in \mathcal{F}^{ij \downarrow}\) is an equilibrium route flow, i.e. \(\mathcal{F}^*(\lambda) = \mathcal{F}^{ij \downarrow}\) for any \(\lambda^i \in (\underline{\lambda}^i, \bar{\lambda}^i)\). Combined with the fact that \(\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\) for any \(\lambda^i \in [\underline{\lambda}^i, \bar{\lambda}^i] \cup [\bar{\lambda}^i, \underline{\lambda}^i]\), we know that \(\mathcal{F}^*(\lambda) \subseteq \mathcal{F}^{ij \downarrow}\) for any \(\lambda\) in regime \(\Lambda_2^\downarrow\).

**Corollary B.1.** If the game \(\Gamma(\lambda)\) has a parallel-route network, then the equilibrium route flow \(f^*\) is unique. Moreover, if there are two populations, then the equilibrium strategy profile is unique in regime \(\Lambda_1^\downarrow\) or \(\Lambda_3^\downarrow\), and can be written as follows:

\[
\text{In regime } \Lambda_1^\downarrow:\quad q_r^1(t^1) = f_r(t^1, \bar{t}^2) - \min_{\bar{t}^2 \in T^2} f_r(t^1, \bar{t}^2), \quad \forall r \in R, \quad \forall t^1 \in T^1, \quad (B.10a)
\]

\[
q_r^2(t^2) = \min_{\bar{t}^1 \in T^1} f_r(\bar{t}^1, t^2), \quad \forall r \in R, \quad \forall t^2 \in T^2, \quad (B.10b)
\]

\[
\text{In regime } \Lambda_3^\downarrow:\quad q_r^1(t^1) = \min_{\bar{t}^1 \in T^1} f_r(t^1, \bar{t}^2), \quad \forall r \in R, \quad \forall t^1 \in T^1, \quad (B.10c)
\]

\[
q_r^2(t^2) = f_r(\bar{t}^1, t^2) - \min_{\bar{t}^2 \in T^2} f_r(\bar{t}^1, \bar{t}^2), \quad \forall r \in R, \quad \forall t^2 \in T^2, \quad (B.10d)
\]

where \((\bar{t}^1, \bar{t}^2)\) is any type profile.

**Proof of Corollary B.1.** Given a parallel route network, we immediately obtain the uniqueness of \(f^*\) from Theorem 1. Then from Proposition 1 any strategy profile that can
induce \( f^* \) can be expressed as in (19). In regime \( \Lambda_{i2} \), we know from Theorem 2 that the constraint \( \{\text{IC}\} \) is tight in equilibrium. Therefore, from (20a) and (20c), we obtain:

\[
\lambda^1 D \overset{20b}{=} \sum_{r \in R} \chi^r \overset{20c}{=} \sum_{r \in R} \max_{t^1, t^2} \left( f_r^*(\hat{t}^1, \hat{t}^2) - f_r^*(t^1, \hat{t}^2) \right) \overset{16}{=} \tilde{J}(f^*) = \lambda^1 D.
\]

Thus, (20c) is tight for any \( r \in R \), i.e.
\[
\chi^r = \max_{t \in T^1} \left( f_r^*(\hat{t}, \hat{t}^2) - f_r^*(t^1, \hat{t}^2) \right).
\]

Additionally, from (20a), \( \chi^2 = \min_{t \in T^1} f_r^*(t^1, \hat{t}^2) \). Thus, \( \chi \) as defined in (20) is unique. Following [Fiacco \& Kyparisis (1986), page 102], the unique equilibrium edge load \( \lambda^1 \) is perturbed in the direction \( z \). Specifically, since the equilibrium edge load \( w^*(\lambda) \) necessarily changes in the direction \( z^i \) in regime \( \Lambda_{i1} \), we have:

\[
\text{[Regime } \Lambda_{i1} \text{]}: \quad \text{Consider any population size vector } \lambda \in \Lambda_{i1}, \text{ there exists a sufficiently small } \epsilon > 0 \text{ such that } \lambda' = \lambda + \epsilon z^j \in \Lambda_{i1}, \text{ i.e. } \lambda' = \lambda + \epsilon > \lambda, \lambda' = \lambda - \epsilon < \lambda, \text{ and the sizes of all other populations remain unchanged. Consider any equilibrium route flow } f^*(\lambda) \in F^*(\lambda) \text{ and any } f^*(\lambda') \in F^*(\lambda'). \]

Then, from Theorem 2, we know from Theorem 2 that constraint \( \{\text{IC}\} \) is tight in equilibrium, and thus \( f^*(\lambda) \) and \( f^*(\lambda') \) satisfy:

\[
\frac{1}{\lambda'} \tilde{J}(f^*(\lambda)) = \lambda' < \lambda^1 = \frac{1}{\lambda'} \tilde{J}(f^*(\lambda')).
\]

Consequently, any equilibrium route flow \( f^*(\lambda) \) for size vector \( \lambda \) is in the feasible domain of (25) for size vector \( \lambda' \), but \( f^*(\lambda) \notin F^*(\lambda') \), because \( \tilde{J}(f^*(\lambda)) = \lambda^1 < \lambda^1' \), i.e. the constraint \( \{\text{IC}\} \) is satisfied, but not tight. Since \( f^*(\lambda') \in F^*(\lambda') \), we must have \( \Psi(\lambda') = \tilde{\Phi}(f^*(\lambda')) < \tilde{\Phi}(f^*(\lambda)) = \Psi(\lambda) \).

Additionally, from (21), we know that \( \Psi(\lambda) = \tilde{\Phi}(w^*(\lambda')) < \tilde{\Phi}(w^*(\lambda)) = \Psi(\lambda) \). Thus, the unique equilibrium edge load \( w^*(\lambda) \) necessarily changes in the direction \( z^ij \) in regime \( \Lambda_{i1} \).

\[
\text{[Regime } \Lambda_{i1} \text{]}: \quad \text{From Theorem 2, } F^*(\lambda) \subseteq F^{ij+1} \text{ for any } \lambda \in \Lambda_{i1}. \text{ Since the equilibrium edge load is unique, we know } w^*(\lambda) = w^{ij+1}. \text{ From (21), we can conclude that } \Psi(\lambda) = \tilde{\Phi}(w^{ij+1}). \text{ Thus, } \Psi(\lambda) \text{ as well as } w^*(\lambda) \text{ remain fixed in regime } \Lambda_{i1}.
\]

\[
\text{[Regime } \Lambda_{i1} \text{]}: \quad \text{Following similar argument in regime } \Lambda_{i1}, \text{ we conclude that } \Psi(\lambda) \text{ monotonically increases in the direction } z^ij \text{ in regime } \Lambda_{i1}. \text{ As a result, } w^*(\lambda) \text{ changes when } \lambda \text{ is perturbed in the direction } z^ij \text{ in regime } \Lambda_{i1}.
\]

Lemma B.2. [Fiacco and Kyparisis (1986), page 102] The value of the potential function in equilibrium, \( \Psi(\lambda) \), is convex with respect to \( \lambda \) if in \( \{\text{OPT-Q}\} \), \( \Phi(q) \) is convex in \( q \), and the constraints are affine in \( q \) and \( \lambda \).

Lemma B.3. [Fiacco (2009), page 3469] If in \( \{\text{OPT-Q}\} \), the objective function \( \Phi(q) \) is convex and continuously differentiable in \( q \), and additionally the set of Lagrange multipliers \( \mu^* \), \( \nu^* \) are nonempty and bounded, then \( \Psi(\lambda) \) is continuous and directionally differentiable in \( \lambda \). Furthermore, for any given \( i, j \in I \), the directional derivative of \( \Psi(\lambda) \) in the direction \( z^ij \) is

\[
\nabla_{z^ij} \Psi(\lambda) = \min_{q* \in Q^*(\lambda)} \max_{(\mu^*, \nu^*) \in (M(q^*), N(q^*))} \nabla \lambda L(q^*, \mu^*, \nu^*, \lambda) z^ij,
\]

where \( M(q^*) \) and \( N(q^*) \) are the sets of Lagrange multipliers \( \mu^* \) and \( \nu^* \) in (12) associated with the BWE \( q^* \in Q^*(\lambda) \).

Proof of Proposition 4. Since the interim belief of population \( j \), \( \beta^j(s, t^j|t^i) \) in (27) is independent with type \( \theta^j \), the equilibrium strategy of the uninformed population \( q^{\theta^i}(t^i) \)
must be identical across all $t^j \in T^j$. Following (15) and (16), the impact of information metric $J^j(q^* \lambda) = \hat{J}^j(f^*) = 0$ for any $q^* \in Q^*(\lambda)$, $f^* \in F^*(\lambda)$ and any $\lambda$. For the sake of contradiction, we assume that the regime $\Lambda_{ij}^3$ is non-empty. From Theorem 2, we know that the constraint (IIC$^j$) must be tight in equilibrium when $\lambda$ is in regime $\Lambda_{ij}^3$. However, since $\hat{J}^j(f^*) = 0$ for any $\lambda$, the constraint (IIC$^j$) is tight only when $\lambda = 1$, i.e. $\lambda = 1 - |\lambda^{-ij}|$. This implies that the regime $\Lambda_{ij}^3$ is indeed empty. Thus, there are at most two regimes $\Lambda_{ij}^1$ and $\Lambda_{ij}^2$. Following Proposition 3, we can conclude that $C^j_\ast(\lambda) \geq C^i_\ast(\lambda)$. □

**Example B.1.** We consider the game with two populations on two parallel routes ($r_1$ and $r_2$) with the following cost functions: $c_1^s(f_1) = f_1 + 15$, $c_1^a(f_1) = 3f_1 + 15$, $c_2(f_2) = 20f_2 + 30$. The prior distribution $\theta$, the total demand $D$, and the information environment are the same as that in Example 1. Although both populations receive the accurate signal of the state with positive probability, we have $\lambda^1 = 1$ as the impact of information on population 2 is zero. Since the free flow travel time on $r_2$ is much higher than that on $r_1$, population 2 travelers exclusively uses $r_1$ regardless of the received signal, see Fig. 5.

**Example B.2.** Consider a game with two populations on two parallel routes ($r_1$ and $r_2$). There are two states: $\{s_1, s_2\}$, each state is realized with probability 0.5. The cost functions are: $c_1^s(f_1) = f_1 + 10$, $c_1^a(f_1) = f_1 + 1$, $c_2^s(f_2) = f_2 + 1$, $c_2^a(f_2) = f_2 + 10$. Population 1 is completely informed, and population 2 is uninformed. The total demand is $D = 1$. We now obtain that $\lambda^1 = \lambda^1 = 1$; thus, regimes $\Lambda_{12}^1$ and $\Lambda_{12}^2$ are empty sets, and population 1 has strictly lower expected cost than population 2 for any $\lambda^1 \in (0, 1)$, see Fig. 6.
C Supplementary material for Section 6

Proof of Proposition 5 Firstly, we prove that for any \( \lambda \in \Lambda^\dagger \), \( F^*(\lambda) \subseteq F^\dagger \). From the definition of \( \Lambda^\dagger \) in (30), we know that for any \( \lambda \in \Lambda^\dagger \), there exists at least one route flow \( f^\dagger \in F^\dagger \) satisfying the constraints in \((\text{OPT-F})\), and hence such \( f^\dagger \) is a feasible solution of the optimization problem \((\text{OPT-F})\); thus \( \hat{\Phi}(f^\dagger) \geq \Psi(\lambda) \). Additionally, since \( f^\dagger \) is an optimal solution of (28), which has the same objective function as \((\text{OPT-F})\) but without the constraints \((\text{IIC})\), we conclude that \( \hat{\Phi}(f^\dagger) \leq \Psi(\lambda) \) for any feasible \( \lambda \) (including \( \lambda \in \Lambda^\dagger \)). Thus, \( \Psi(\lambda) = \hat{\Phi}(f^\dagger) \), and \( f^\dagger \) is an equilibrium route flow. Analogous to the argument in proof of Theorem 1, the equilibrium edge load equals to \( w^\dagger \). Since the set \( F^\dagger \) in (29) contains all route flows such that the induced edge load is \( w^\dagger \), we can conclude that the set of equilibrium route flow \( F^*(\lambda) \subseteq F^\dagger \) for any \( \lambda \in \Lambda^\dagger \).

Next, we prove that \( \Lambda^\dagger = \arg \min_\lambda \Psi(\lambda) \). We have argued in the first part of the proof that for any \( f^\dagger \in F^\dagger \), \( \hat{\Phi}(f^\dagger) \leq \Psi(\lambda) \) for any feasible \( \lambda \); and since for any \( \lambda \in \Lambda^\dagger \), \( F^*(\lambda) \subseteq F^\dagger \), we have \( \Psi(\lambda) = \hat{\Phi}(f^\dagger) \). Therefore, \( \hat{\Phi}(f^\dagger) = \min_\lambda \Psi(\lambda) \), and \( \Lambda^\dagger \subseteq \arg \min_\lambda \Psi(\lambda) \). Additionally, for any \( \lambda \in \arg \min_\lambda \Psi(\lambda) \), we have \( \Psi(\lambda) = \min_\lambda \Psi(\lambda) = \hat{\Phi}(f^\dagger) \). Since \( F^\dagger \) includes all route flows that satisfy (14a)-(14c) and attain the minimum value of \( \Psi(\lambda) \), any equilibrium route flow \( f^* \in F^*(\lambda) \) for \( \lambda \in \arg \min_\lambda \Psi(\lambda) \) must be in \( F^\dagger \). Hence, such \( \lambda \) must also be in \( \Lambda^\dagger \) defined in (30), i.e. \( \arg \min_\lambda \Psi(\lambda) \subseteq \Lambda^\dagger \). We can therefore conclude that \( \Lambda^\dagger = \arg \min_\lambda \Psi(\lambda) \).

From Lemma 5, we know that the function \( \Psi(\lambda) \) is convex in \( \lambda \). Additionally, the set of feasible population size vector \( \lambda \) is a closed convex set. Consequently, the set \( \Lambda^\dagger = \arg \min_\lambda \Psi(\lambda) \) is convex and non-empty.

Finally, we show that \( w^*(\lambda) = w^\dagger \) if and only if \( \lambda \in \Lambda^\dagger \). From the first part of the proof, we know that \( F^*(\lambda) \subseteq F^\dagger \). Therefore, the unique equilibrium edge load is \( w^\dagger \), which does not depend on \( \lambda \). Additionally, for any feasible \( \lambda \notin \Lambda^\dagger \), from the second part of the proof, we know that \( \Psi(\lambda) > \hat{\Phi}(w^\dagger) \). Thus, \( w^*(\lambda) \neq w^\dagger \).
D Extension to Networks with Multiple Origin-destination Pairs

In this section, we extend our model to networks with multiple origin-destination pairs, and show that all the results presented in the paper still hold. Consider a network with a set of origin-destination (o-d) pairs $\mathcal{K}$. Each o-d pair $k \in \mathcal{K}$ is connected by the set of routes $\mathcal{R}_k$. The set of all routes is $\mathcal{R} = \bigcup_{k \in \mathcal{K}} \mathcal{R}_k$. The demand of travelers between o-d pair $k \in \mathcal{K}$ is $D_k > 0$. The information environment – state, TISs, signals and common prior – is the same as that introduced in Sec. 3.1. The fraction of travelers between o-d pair $k \in \mathcal{K}$ who subscribe to TIS $i \in \mathcal{I}$ is $\lambda^i_k$. A feasible size vector $\lambda = (\lambda^i_k)_{k \in \mathcal{K}, i \in \mathcal{I}}$ satisfies $\lambda^i_k \geq 0$ for any $k \in \mathcal{K}$ and any $i \in \mathcal{I}$, and $\sum_{i \in \mathcal{I}} \lambda^i_k = 1$ for any $k \in \mathcal{K}$. We denote the strategy profile as $q^i = (q^{i}_{r,k}(t^i))_{r \in \mathcal{R}_k, i \in \mathcal{I}, t^i \in \mathcal{T}^i, k \in \mathcal{K}}$, where $q^{i}_{r,k}(t^i)$ is the amount of travelers in population $i$ who take route $r$ between o-d pair $k$ when the signal is $t^i$. A strategy profile $q$ is feasible if it satisfies:

$$
\sum_{r \in \mathcal{R}_k} q^{i}_{r,k}(t^i) = \lambda^i_k \cdot D_k, \quad \forall i \in \mathcal{I}, \quad \forall t^i \in \mathcal{T}^i, \quad \forall k \in \mathcal{K},
$$

$$
q^{i}_{r,k}(t^i) \geq 0, \quad \forall r \in \mathcal{R}_k, \quad \forall i \in \mathcal{I}, \quad \forall t^i \in \mathcal{T}^i, \quad \forall k \in \mathcal{K}.
$$

For any feasible strategy profile $q$, the induced route flow vector is $f^i = (f_{r,k}(t))_{r \in \mathcal{R}_k, k \in \mathcal{K}, t^i \in \mathcal{T}}$, where $f_{r,k}(t)$ is the flow on route $r$ induced by travelers between o-d pair $k$ when the type profile is $t$:

$$
f_{r,k}(t) = \sum_{i \in \mathcal{I}} q^{i}_{r,k}(t^i), \quad \forall r \in \mathcal{R}_k, \quad \forall k \in \mathcal{K}, \quad \forall t \in \mathcal{T}.
$$

(D.11)

The aggregate edge load on edge $e \in \mathcal{E}$ when the type profile is $t$ can be written as follows:

$$
w_e(t) = \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \sum_{r \in \{\mathcal{R}_k \mid r \in \mathcal{E}\}} q^{i}_{r,k}(t^i) = \sum_{k \in \mathcal{K}} \sum_{r \in \{\mathcal{R}_k \mid r \in \mathcal{E}\}} f_{r,k}(t).
$$

(D.12)

The expected cost of each route $r \in \mathcal{R}$ given type $t^i \in \mathcal{T}^i$, denoted $\mathbb{E}[c_r(t^i)|t^i]$, can be written as in (5). A feasible strategy profile $q^*$ is a BWE if for any $k \in \mathcal{K}$, any $i \in \mathcal{I}$, and any $t^i \in \mathcal{T}^i$:

$$
\forall r \in \mathcal{R}_k, \quad q^{i*}_{r,k}(t^i) > 0 \Rightarrow \mathbb{E}[c_r(q^*)|t^i] \leq \mathbb{E}[c_{r'}(q^*)|t^i], \quad \forall r' \in \mathcal{R}_k.
$$

We now state the extensions of our results to the network with $\mathcal{K}$ o-d pairs. Firstly, we can check that the following function of $q$ is a weighted potential function of the Bayesian congestion game with $\mathcal{K}$ o-d pairs:

$$
\Phi(q) = \sum_{e \in \mathcal{E}} \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{T}} \pi(s, t) \int_0^{\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \sum_{r \in \{\mathcal{R}_k \mid r \in \mathcal{E}\}} q^{i}_{r,k}(t)} c^s_e(z) dz,
$$

Therefore, given any size vector $\lambda$, the set of equilibrium strategy profiles $\mathcal{Q}^*(\lambda)$ can be solved by the same convex optimization problem (OPT-Q) in Theorem 1. The equilibrium edge load $w^*(\lambda)$ is unique.
Secondly, with simple modifications, we characterize the set of feasible route flow vectors \( F \) as follows:

\[
F \triangleq \{ f \mid \forall k \in \mathcal{K}, (f_{r,k}(t))_{r \in \mathcal{R}_k, t \in \mathcal{T}} \text{ satisfies (14a) – (14d)} \},
\]

and the set of equilibrium route flows \( F^*(\lambda) \) can be solved by the optimization problem \((\text{OPT-F})\) in Proposition 2. Particularly, the information impact constraint (14d) for o-d pair \( k \in \mathcal{K} \) and population \( i \in \mathcal{I} \) now becomes:

\[
D_k - \sum_{r \in \mathcal{R}_k} \min_{t^i \in \mathcal{T}^i} f_{r,k}(t^i, t^{-i}) \leq \lambda_i^k D_k, \quad \forall t^{-i} \in \mathcal{T}^{-i}, \quad \forall i \in \mathcal{I}, \quad \forall k \in \mathcal{K}. \quad \text{(D.13)}
\]

That is, the impact of information sent by TIS \( i \) on the route flows between o-d pair \( k \) is bounded by the amount of travelers who subscribe to TIS \( i \) and travel between o-d pair \( k \).

Thirdly, for any o-d pair \( k \) and any pair of TISs \( i \) and \( j \), we can analogously analyze how the equilibrium outcomes change with the sizes \( \lambda_i^k \) and \( \lambda_j^k \), while the sizes of all other populations remain unchanged. We denote this direction of perturbing \( \lambda \) as \( z_{ij}^k \). We can show that Theorem (2) holds: three regimes (one or two may be empty) can be distinguished by whether or not the information impact all the travelers between o-d pair \( k \) who subscribe to TIS \( i \) (resp. \( j \)), i.e. whether or not (D.13) is tight at the optimum of \((\text{OPT-F})\).

Fourthly, the relative value of information, denoted \( V_{ij}^{k^*}(\lambda) \), is the travel cost saving experienced by travelers between o-d pair \( k \) who subscribe to TIS \( i \) compared with travelers who subscribe to TIS \( j \). Analogous to Lemma 5 we obtain:

\[
V_{ij}^{k^*}(\lambda) = C_{ij}^{k^*}(\lambda) - C_{ij}^{k^*}(\lambda) = -\frac{1}{D_k} \nabla \Phi_{z_{ij}^k}(q).
\]

By applying sensitivity analysis, we can show that \( V_{ij}^{k^*}(\lambda) \) is positive, zero, and negative in the three regimes respectively.

Finally, we analyze how travelers between each o-d pair \( k \in \mathcal{K} \) choose TIS subscription. By dropping the information impact constraints (D.13) in \((\text{OPT-F})\) and following (29) – (30), we obtain the extension of Theorem 4: characterization of the set \( \Lambda^f \), which is the set of size vectors induced by travelers’ choice of TISs. We note that how an individual traveler between o-d pair \( k \) chooses information subscription not only depends on the choices of other travelers between the same o-d pair, but also the choices of travelers who travel between other o-d pairs. This is because travelers between different o-d pairs may take common routes and edges, and hence impact each others costs.
This figure "diagram_map.png" is available in "png" format from:

http://arxiv.org/ps/1808.10590v3