A General Equilibrium Model for Integrated CAV Ridesourcing and Transit Services for the Morning Commute

Rong Fan

Department of Civil and Environmental Engineering University of Washington, Seattle, WA, 98195 Email: rongfan6@uw.edu

Dan MacCabe

Department of Civil and Environmental Engineering University of Washington, Seattle, WA, 98195 Email: dmccabe@uw.edu

Xuegang (Jeff) Ban (Corresponding Author)

Department of Civil and Environmental Engineering University of Washington, Seattle, WA, 98195 Email: banx@uw.edu

Word Count: 5849 words + 6 table (250 words per table) = 7349 words

Submitted [August 01, 2020]

ABSTRACT

Commuting congestion increases along side the prosperity of urban cities. With the rapid development of ridesourcing services and the advances of the connected and automated vehicles (CAV), researchers are seeking innovative approaches to alleviate commuting congestion by integrating CAVbased ridesourcing and transit services. We propose a general equilibrium model for an integrated, multimodal CAV ridesourcing and transit system. Our model captures the economic behaviors and interactions of the major players (i.e.. the ridesourcing company and customers) in the commuting problem by optimizing the profit of the ridesourcing company and the utility of customers, as well as considering the network congestion. Results show that the demand for shared rides and transit are affected by the relative costs of different types of travel modes of the integrated system. While transit uses generally reduces congestion, ridesharing alone may still cause higher congestion compared with solo driving because of the deadhead miles. Our model can systematically investigate the mode choices of customers and measure the resulting congestion effect in a multimodal network, which helps bring valuable insights to transportation planers, transit agencies, and ridesourcing companies.

Keywords: Integrated Transit, Ridesourcing, Connected and Automated Vehicles, General Equilibrium Model

1. INTRODUCTION

Enabled by increasing smartphone uses, ridesourcing companies (e.g., Uber/Lyft/Didi) have made remarkable progress on serving urban travel demands through real time matching of drivers and customers (*[1](#page-16-0)*). In New York City, e.g., ridesourcing served 4.2 billion trips in 2018, up from about 2 billion trips in 2016, which was also more than eight times of the trips served by taxis in 2018 (*[2](#page-16-1)*). Platforms that offer "ridesourcing" (also referred to as "e-hailing" or "ride-hailing" (*[3](#page-16-2)*)) may also provide shared-ride (i.e. "ridesplitting") services for customers from nearby locations (mostly two), such as Uber Pool. In this paper, we use "single rides" or "shared rides" to distinguish ridesourcing services for single customer pickups or multiple customer pickups, respectively (*[4](#page-16-3)*).

Meanwhile, rapidly evolving vehicle automation technology has the potential to reduce vehicle ownership and increase the need for ridesourcing services, which is expected to trigger another revolution in urban mobility (*[5](#page-16-4)*). However, the rapid growth of ridesourcing may inevitably lead to more car travels and urban congestion (*[6,](#page-16-5) [7](#page-16-6)*). To mitigate such negative impacts, ridesourcing, especially when CAVs are widely deployed, has to be properly integrated with public transit in order to provide more accessible and efficient means to all travelers in urban areas.

On-demand transit system, the integration of demand responsive services and transit, has been studied and practiced long before the era of ridesourcing. For example, Stein studied the integration of Dial a Ride services with fixed transit routes to solve the first/last mile problem (*[8](#page-16-7)*), which has been the main challenge for many potential transit users. Later on, studies on similar problems showed that the integration of on-demand or flexible route services with fixed route transit could reduce the operation cost of transit agencies (*[9,](#page-16-8) [10](#page-16-9)*).

The wide deployment of ridesourcing services has inspired research on the integration of ridesourcing and transit (*[11](#page-16-10)*). Chen et al. compared the line-based design and the zone-based design for the integration of ridesourcing and transit (*[12](#page-16-11)*). Their analytical and simulation results both showed that line-based services can achieve higher efficiency and allow ridesharing. Ma et al. proposed a ridesharing scheme with integrated transit in which ridesourcing serves either the whole trip or the first/last mile of a transit route (*[13](#page-16-12)*). Their objective was to optimize vehicle dispatch and idle vehicles relocation for an integrated, multimodal transit-rideshare system. The numerical experiments indicated that the transitrideshare system outperforms the rideshare-only system by 32% reduction in user travel time and 64% reduction in vehicle travel time. Compared with rideshare-only system, the transit-ridesharing system also reduces the operation cost and customer waiting times, and achieves better performance in regions where passenger demands are heterogeneous. Pinto et al. proposed a bi-level mathematical programming formulation for the joint transit network redesign and mobility service fleet size determination (*[14](#page-17-0)*). The integration in their study was to replace inefficient transit routes/patterns with shared autonomous mobility services, while the on-demand first/last mile for transit is not considered. In addition, their model only had one objective function to minimize the disutility of customers, with no platform profit maximization or congestion consideration. Ban et al. proposed a general equilibrium model that consists of the optimization of three main players: service providers, passengers and the network congestion (*[6](#page-16-5)*). Their study provided insights for the transportation planning agencies on the congestion impact of ridesourcing. The deadhead miles of either taxi or ridesourcing services generally exacerbate the network congestion, but when the demand pattern has high level of symmetry, the deadhead miles can be significantly reduced.

In addition to studies on the operation of the integrated system of transit and ridesourcing, researchers also investigated the user experience and preferences of passengers. Yan et al. evaluated travelers' responses to the integrated transit pilot MTransit and showed that ridesourcing could complement transit by serving the first/last mile and/or replacing fixed transit routes with low usage, leading to reduced passenger waiting times and lower operation costs of transit agencies (*[15](#page-17-1)*).

It suffices to say that most studies so far have focused on the integrated ridesourcing and transit mode without considering other related modes that customers may choose, e.g., ridesourcing services alone (eithr single rides or shared rides) for an entire trip. Pinto et al. did consider multimodal whereas platform profit maximization or congestion effect was not considered (*[14](#page-17-0)*).

In this paper, we aim to develop a modeling framework that is not only mathematically rigorous, but also captures the key behaviors and interactions of the major players when integrating ridesourcing with transit. To illustrate our approach, we focus on the morning commute, one of the most important daily trips, which also experiences the most congestion. We model fixed-route mass transit, such as light rail, subways, or bus rapid transit (BRT) on fixed schedules. As a result, there are three major players when integrating ridesourcing and transit services with CAVs: i) Customers who choose mobility services based on their values of time and other personal/social characteristics; ii) a ridesourcing platform providing both single ride and shared ride services, which usually serves customers and dispatches vehicles (CAVs) to maximize its profit; and iii) the route choices of CAVs that impact the overall network congestion. Clearly, the three players have distinct objectives; however, they interact with each other on the multimodal transportation network of an urban area, resulting in the use pattern of each travel mode and the overall network congestion. Therefore, understanding and modeling the behavior and the interactions of these three players are crucial to better integrate ridesourcing and transit and to evaluate their collective effect on the urban transportation system.

A general equilibrium model on an integrated multimodal network is proposed to capture the behavior and interactions of the three players, which is sensitive to customers' value of time and system congestion effects. Our model significantly extends the model in Ban et al. (*[6](#page-16-5)*) to encompass shared rides and transit services under the CAV environment. The main contributions of this paper are:

i) Develop methods to model the behavior of the three major players (customers, ridesourcing providers, CAV route choices) on a multimodal network, including the development of an extended network structure;

ii) Formulate a general equilibrium model with single rides, shared rides, and integrated CAV ridesourcing and transit services, for which formal analysis can also be conducted including the existence and uniqueness of its solution, and solution methods;

iii) Evaluate the network effect as a result of different CAV dispatching strategies, different values of time of customers, and different usage rates of transit. This may provide useful insights on developing policies to manage ridesourcing and to better integrate it with transit services.

2. METHODS

2.1 Problem statement

Figure 1 (a) illustrates the morning commute scenario. There are four types of modes, $m \in$ {1,2,3,4}, which can serve commuters from their residential areas to the worksites: i) single rides from origin to destination, $m=1$; ii) shared rides that take passengers from two separate nearby locations from origin to destination, $m=2$. iii) single rides as the first mile then transfer to transit, $m=3$; iv) shared rides that take passengers from two separate locations as the first mile then transfer to transit, m=4. Single rides services and shared rides services are operated by the same ridesourcing platform. For simplicity, we refer to the four types of modes as taxi (m=1), ridesharing (m=2), Ttransit (m=3), RStransit (m=4); see **Figure 1 (b)**. A customer chooses a particular mode based on his/her value of time and other characteristics.

	Single ride	Shared ride
Ridesourcing mode Taxi, m= 1		Ridesharing, $m=2$
Integrated mode		Ttransit, $m=3$ RStransit, $m=4$

Figure 1 Integrated multimodal network

Here we do not consider the mode where a commuter owns a CAV and drives alone to the worksite (i.e. the "solo driving" mode as defined in (*[6](#page-16-5)*)), in order to better investigate the interaction between CAV ridesourcing and transit. Also the transit services considered here are mass transit with fixed routes/schedules (i.e. on-demand transit services are not considered), such as light rail, subway, or BRT, which usually run on separate right-of-way and thus have little or no contribution to vehicular traffic congestion. The proposed modeling framework may be extended to include solo driving, ondemand transit, and street bus services, which we leave for future research. Also in practice, according to Li's et al. (16), more than 90% of shared rides occur for two pickup locations. For the current study, therefore, we only consider shared rides in which a CAV picks up customers from (up to) two separate nearby locations. This is also consistent with the ridesplitting services in most ridesourcing platforms (e.g. Uber Pool).

2.2 General equilibrium overview

We propose a general equilibrium model with three modules (**Figure 2**). In the ridesourcing choice module, we maximize the revenue of the provider, the endogenous variable is the CAV dispatch (i.e. vehicle supply), while customer demand and route choice are exogenous variables. In the customer choice module, we minimize the disutility of customers, i.e., customers mode choices that provide them the highest utility. The decision variables are the demand of each mode with vehicle supply and route choices are the exogeneous variables. The network congestion module captures the flow interaction and congestion effect due to the choices and interaction of customers and service providers. We model the behavior of vehicles' route choices according to Wardrop's first principle, i.e., the CAVs always choose the route with the minimum travel time. The route choice is the decision variable, while demand and vehicle supply are exogenous variables. The three modules combine to form a general equilibrium model.

Figure 2 A summery of the general equilibrium model

2.3 Extended network structure

We construct an extended network of two layers to model the integared multimodal network, similar to but more concise compared to the method in Di and Ban (*[16](#page-17-2)*). **Figure 3** shows the extended structure for a small network. **Figure 3 (b)** is the ridesourcing layer, where node 5 is the vehicle destination for ridesourcing $(m=1,2)$, and transit station node 4 is the vehicle destination for the ingegrated modes (m=3,4). For the transit layer (**Figure 3 (a)**), we define transit nodes 6 and 7, where node 6 has the same location with node 4 in the ridesourcing layer, and node 7 has the same location as node 5 (**Figure 3 (c)**). Vehicle flow is based on the ridesourcing layer, passenger flow is based on both layers. Recall that our model assumes that transit runs on separate right-of-way (e.g., subway or light rail; express routes for buses) and therefore does not interact with the ridesourcing layer when congestion is concerned. To simplify the discussion, in this paper, we only consider CAV ridesourcing serving the first mile of transit and thus assume customers' destinations coincide with transit stations. Similar method can be applied to model the scenario where CAV ridesouring serves the last mile of transit when customers' destinations are different from transit stations. The proposed model can also deal with multiple transit stations at the origins and/or at the destinations, although only one such station is shown (at the origins or at the destinations) in **Figure 3** for illustration purposes.

Figure 3 Extended network structure of the small network

3 MODEL FORMULATION

We define the following sets in our model: $m \in \{1,2,3,4\}$, four types of modes, introduced in section 2.1; N, the set of nodes; O, the set of origins, where vehicles pick up passengers; D^{ν} , the set of destinations of ridesourcing vehicles, where vehicles drop off all passengers and then head to the next pick-up location; D^c , the set of destinations of customers, which are the worksites; $K \in O \times D^c$, the set of origin-destination (OD) pairs for customers; T , the set of transit stations, where customers of mode 3 or 4 transfer to transit. In the model formulation, O_k denotes the origin of an OD pair k , D_k^{ν} denotes the vehicle destination of the OD pair *k*, D_k^c denotes the customer destination of the OD pair *k*, T_k denotes the transit station of the OD pair *k* for mode 3 or 4. That is, vehicles and customers have different destinations for the integrated transit modes (m=3,4), but have the same destination if the whole trip is served by the ridesourcing modes (m=1,2). Here we define a mapping set $MAP := \{(m, D^{\nu}, D^c), k \in K | (1, D_k^c, D_k^c),$ $(2, D_k^c, D_k^c)$, $(3, D_k^v, D_k^c)$, $(4, D_k^v, D_k^c)$ to connect the sets of customer destinations and CAV destinations of each mode. To illustrate these sets using the network in **Figure 3**, we have $N = \{1,2,3,4,5\}$, O $:=$ {1,2,3}, $D^v :=$ {4,5}, $D^c :=$ {5}, $T_k :=$ {6}, $MAP :=$ { (m, D^v, D^c) | (1,5,5), (2,5,5), (3,4,5), (4,4,5)}.

3.1 Ridesourcing choice module (*Module I***)**

The revenue of each shared-ride vehicle consists of the fixed fare of each passenger and the timeand distance-based fares. The fixed fare $F_{O_k}^m$ and $F_{O_{k'}}^m$, are based on the locations of the origin nodes, O_k and O_{kt} . Notice that although the vehicle takes detours to pick up passengers, the fare that passengers need to pay is based on the distance between their origins and the destination. Thus the time- and distance-based fares are formulated as $a_1^m(t_{O_kO_k}+t_{O_kO_k})$ and $a_2^m(d_{O_kO_k}+d_{O_{k}/O_k})$, where a_1^m is the time-based fare rate, and α_2^m is the distance based fare rate. Notice here that for simplicity, we assume the two groups of passengers on a shared ride go to the same destination (i.e. the same transit stop or the same downtown worksite). The modeling method proposed here can be properly revised to incorporate the case when the two destinations are separate and close to each other.

The cost of each shared-ride vehicle also includes time- and distance-based cost of the actual route taken, represented by $\beta_1^m(t_{j\omega_k}+t_{\omega_k\omega_{k'}}+t_{\omega_{k'}\omega_k})$ and $\beta_2^m(d_{j\omega_k}+d_{\omega_k\omega_{k'}}+d_{\omega_{k'}\omega_k})$, where β_1^m is the time-based cost rate and β_2^m is the distance-based cost rate. Since we specifically model CAVs, the time-based cost will be lower than it is for traditional taxis. However, if a vehicle takes longer than usual to finish a trip, it indicates less efficiency and lower customer satisfaction, which cause extra cost to the platform. In summary, the profit function of a shared-ride vehicle that drops off the previous passenger(s) at *j* and then picks up passengers first at O_k and then at O_k can be formulated as:

$$
R_{jkk'}^m = F_{O_k}^m + F_{O_{kl}}^m + \underbrace{\alpha_1^m (t_{O_k D_k} + t_{O_{kl} D_k})}_{time\ based} + \underbrace{\alpha_2^m (d_{O_k D_k} + d_{O_{kl} D_k})}_{distance\ based} - \underbrace{\beta_1^m (t_{jO_k} + t_{O_k O_{kl}} + t_{O_{kl} D_k})}_{travel\ time\ based\ cost} - \underbrace{\beta_1^m (t_{jO_k} + t_{O_k O_{kl}} + t_{O_{kl} D_k})}_{travel\ distance\ based\ cost}
$$

Similarly, the profit of a single-ride vehicle that drops off the previous passenger(s) at j and then picks up passengers at O_k and drop off passengers at destination D_k can be formulated as:

$$
R_{jk}^{m} = F_{O_k}^{m} + \underbrace{\alpha_1^{m} t_{O_k D_k}}_{time\ based} + \underbrace{\alpha_2^{m} d_{O_k D_k}}_{distance\ based} - \underbrace{\beta_1^{m} (t_{jO_k} + t_{O_k D_k})}_{travel\ time\ based\ cost} - \underbrace{\beta_2^{m} (d_{jO_k} + d_{O_k D_k})}_{travel\ distance\ based\ cost} \forall m = 1,3
$$

For $m = 1$ or 2, commuters take single rides or shared rides to their actual destinations (e.g., in **Figure 3**, the drop-off location is node $j=5$; for $m = 2$ or 4, commuters take single rides or shared rides to the transit stations (not their actual destinations), e.g., in **Figure 3**, the drop-off location is node $j=4$ (the same as node 6 in the transit layer). This needs to be properly modeled.

The decision variables of the ridesourcing choice module are z_{jkk}^m , (where O_k , O_{k} , \in 0) for shared-ride vehicles and z_{jk}^m (where $O_k \in O$) for single-ride vehicles. z_{jkk}^m , denotes the number of sharedride vehicles currently at destination *j* and will next pick up customers first at O_k and then at $O_{k'}$. z_{jk}^m denotes the number of single-ride vehicles currently at destination j and will next pick up customers at O_k . Although there are four different types of modes, the ridesourcing platform is only providing two types of services, either single rides or shared rides. Here we assume that for the same service, the platform has the same price strategies for the first-mile trips and whole trips. The objective function of *Module I* is to maximize the total profit from single-ride and shared-ride services:

$$
z_{jk}^{m} \geq 0, z_{jkk'}^{m} \geq 0, q_{kk'}^{m}, b_{kk'}^{m} \underbrace{\sum_{m=1,3} R_{jk}^{m} z_{jk}^{m}}_{single\ rate} + \underbrace{\sum_{m=2,4} (R_{jkk'}^{m} z_{jkk'}^{m} - \omega^{m} \cdot |q_{kk'}^{m}| - \omega^{m} \cdot |b_{kk'}^{m}|)}_{shared\ rate}
$$

Subject to:

Constraints for single rides, $m = 1,3$: $\sum_{j\in D} z_{jk}^m \geq Q_k^m$ $\forall k \in K$ vehicle supply satisfies customer demand $\sum_{k \in K} z_{jk}^{m} = \sum_{k': j = D_k} Q_{k'}^{m}$ $\forall j \in D$ vacant single-ride CAVs are avaible again to serve next trip

Constraints for shared rides, $m = 2,4$:

 $\sum_{j\in D} z_{jkk\prime}^m - Q_{k,kk'}^m = b_{kk'}^{+m} - b_{kk'}^{+m} \quad \forall k \in K \quad shared \text{ vehicles pick up 2 customers in each trip}$ $\sum_{j\in D}z_{jkk\prime}^m-Q_{k\prime,kk\prime}^m=q_{kk'}^{+m}-q_{kk'}^{+m} \quad \forall k\in K \quad shared\ vehicles\ pick\ up\ 2\ customers\ in\ each\ trip$ $\sum_{k,k} \sum_{j \in K} z_{jkk}^m = \sum_{k \in K} \sum_{k',j} z_{k,j}^m$, $Q_{k',kk'}^m$ $\forall j \in D$ vacant shraed-ride CAVs are avaible again

Constraint for the total number of vehicles:

\n
$$
\sum_{m=1,3} \left(\sum_{k \in K} \sum_{j \in D} z_{jk}^{m} t_{jO_k} + \sum_{k \in K} Q_{k}^{m} t_{O_k D_k} \right) + \n\frac{\sum_{m=2,4} \left(\sum_{k,k' \in K} \sum_{j \in D} z_{jkk'}^{m} t_{jO_k} + \sum_{k,k' \in K} Q_{k,kk'}^{m} t_{O_k O_{k'}} + \sum_{k,k' \in K} Q_{k,kk'}^{m} t_{O_{k} D_{k'}} \right)}{\n\text{number of vehicles for shared ride}}
$$
\nSince the probability function $\sum_{k=1}^{m} \sum_{k=1}^{m} \sum_{k=1}^{m} z_{jkl}^{m} t_{jO_k} + \sum_{k,k' \in K} Q_{k,kk'}^{m} t_{O_{k} D_{k'}} = \sum_{k=1}^{m} \sum$

Here Q_k^m is the demand of mode m along OD pair k. $Q_{kk'}^m$ and $Q_{k'k}^m$ are the demand of the fist pickups and second pickups of the shared ride services (m=3,4). As illustrated in **Figure 2**, customer demand variables are exogoneous to *Module I*.

A mismatch occurs when shared rides fail to pick up 2 passengers from 2 different origin nodes. $q_{kk'}^m$ and $b_{kk'}^m$ are defined to capture the mismatch of shared rides. There is no mismatch when both $q_{kk'}^m$ and $b_{kk'}^m$ equals 0. ω^m is the rate of penalty, higher value means higher cost per mismatch. Constraints for *Module I* ensure that: i) the number of vehicles dropping off customers at destination \dot{j} is equal to the number of vehicles departing from j ; ii) each customer request is served; iii) shared-ride vehicles can only pick up at most two passengers from up to two different origins, and penalization will be imposed if they only pack up passengers from only one location; iv) the number of vehicles in operation is no larger than the total number of vehicles operated by the platform.

3.2 Customer choice module (*Module II***)**

Customers can choose to take rideshourcing to a transit station or directly to worksites. Thus, the customer flow is present on both the ridesourcing layer and the transit layer. When calculating the disutility of the shared-ride customers, either for the first mile or for the whole trip, there is a trade-off between precisely capturing the behavior of the customers and the feasibility of the model. For shared-ride trips, the first customer tends to have a longer ride time while the second customer tends to have a longer waiting time. But if the shared-ride trips benefit the first customer more than the second customer, it will be hard for the model to find a second customer so that the model is feasible. Thus, we simplify travel/waiting timebased disutility so that two customers in the same shared-ride trip expect the same (and the worst case) disutility caused by the detour. The disutility of a shared-ride customer that is picked up first can be formulated as,

$$
V_{kk'}^m = F_{O_k}^m + \underbrace{\alpha_1^m t_{O_k D_k}}_{fixed} + \underbrace{\alpha_2^m d_{O_k D_k}}_{fixed\text{fixed}} + \underbrace{\alpha_2^m d_{O_k D_k}}_{distance\text{ based}} + \underbrace{\gamma_1^m (t_{O_k O_{k'}} + t_{O_k D_k})}_{travel\text{time}} + \underbrace{\gamma_2^m t_{O_k O_{k'}} + w_{kk'}}_{disutility\text{ due}} + \underbrace{\gamma_3^m \lambda_{kk'}^m}_{disutility\text{ due}} + \underbrace{\gamma_3^m \lambda_{kk'}^m}_{to\text{parallelity\text{ due}}}
$$

Where O_k is the first pick-up location, O_k , is the second pick-up location. We have disutility due to the fare of the shared rides. $F_{O_k}^m$ is the location-based fixed fare. The time-based fare $\alpha_1^m t_{O_k D_k}$ and the distance-based fare $\alpha_2^m d_{O_k D_k}$ depend only on the OD pair $O_k \to D_k$, which is reasonable because the additional disutility of shared rides is already considered in other terms. γ_1^m is the disutility rate for travel time. The travel time based disutility consists of the travel time of OD pair $O_k \to D_k$ and the travel time of detour $O_k \to O_k$. γ_2^m is the waiting time disutility rate, $w_{kk'}^m$ is the waiting time based disutility from the previous drop off location to the first pick-up location. Thus the summation of $\gamma_2^m t_{O_k O_{k'}}$ and $w_{kk'}^m$ captures the worest case waiting time, from the vehicle's previous drop-off location *j* to the second shared-ride customer. The matching of the two customers in a shared ride is also regarded as the cost of the platform, which is modelled as the Lagrangian multiplier of the shared ride demand constriants in *Module I*, denoted as $\gamma_3^m \lambda_{kk}^m$, where γ_3^m is the rate of matching cost and λ_{kk}^m , is the multiplier. See [\(6\)](#page-16-5) for more discussions of why this multiplier could be regarded as (proportional to) the matching cost. The diutility of the second picked-up customer, V_{krk}^m , can be derived similarly, which is omitted here.

We assume high capacity for the transit route $T_k \rightarrow D_k$, thus the travel time and distance are constant for the transit route. The travel time and travel distance can be represented by the same parameter, $t_{T_k D_k} \cdot const = d_{T_k D_k}$. The disutility of taking transit consists of the time/distance-based transit fare $\alpha_3^m d_{T_k D_k}$, time/distance-based disutility $\gamma_4^m d_{T_k D_k}$ and the transfer cost γ_5^m . When $m = 1,2$, ridesourcing vehicles serve the whole trip, so $Tr_k^m = 0$. Therefore the the disutility of transit is

$$
Tr_{k}^{m} = 0 \quad \forall m = 1,2
$$

\n
$$
Tr_{k}^{m} = \underset{\text{travel distance}}{\underbrace{\alpha_{3}^{m}d_{T_{k}D_{k}}}} + \underset{\text{travel distance}}{\underbrace{\gamma_{4}^{m}d_{T_{k}D_{k}}}} + \underset{\text{disutility of}}{\underbrace{\gamma_{K}^{m}D_{k}}}
$$

\n
$$
\forall m = 3,4
$$

\n
$$
travel \overline{distance} \quad \text{based distributive} \quad \text{distributive}
$$

Similar to shared rides, the disutility of a customer of a single ride trip can be formulated as,

$$
V_k^m = F_{O_k}^m + \underbrace{\alpha_1^m t_{O_k D_k}}_{fixed\text{ fare}} + \underbrace{\alpha_1^m t_{O_k D_k}}_{distance\text{ based}} + \underbrace{\gamma_1^m t_{O_k D_k}}_{fixed\text{ trace} time} + \underbrace{\gamma_1^m t_{O_k D_k}}_{travel\text{ time}} + \underbrace{\gamma_1^m}_{distility\text{ due}} + Tr_k^m \quad \forall m = 1,3
$$

Customer waiting time is calculated as the average deadhead miles from the previous destinations to the next origin (first origin for shared ride trips); see (*[6](#page-16-5)*),

$$
w_k^m = \gamma_2^m \frac{\sum_{j \in D} z_{jk}^m t_{j0_k}}{\sum_{j \in D} z_{jk}^m} \quad \forall k \in K, m = 1,3
$$

$$
w_{kk'}^m = \gamma_2^m \frac{\sum_{j \in D} z_{jkk'}^m t_{j0_k}}{\sum_{j \in D} z_{jkk'}^m} \quad \forall k, k' \in K, m = 2,4
$$

The decision variable for *Module II* is the customer demand. The customer demand for single rides is Q_k^m . The customer demand for shared rides are $Q_{kk'}^m$ and $Q_{k'k}^m$ for the first pick-up and second pick-up, respectively. The objective function minimizes the disutility of customers:

min $q_{k'k'}^{m}q_{kk'}^{m}$, q_{k}^{m} $\sum_{m=1,3} V_k^m Q_k^m + \sum_{m=2,4} V_{kk}^m Q_{kk}^m + \sum_{m=2,4} V_{k'k}^m Q_{k'kk}^m$ Subject to: $\sum_{m=1}^{4} Q_{k}^{m} = Q_{k} \quad \forall k \in K$ $\sum_{k' \in K} (Q_{k'k}^m + Q_{kk'}^m) = Q_k^m$ $\forall m = 2, 4, k \in K$

The constraints for *Module II* ensure that: i) the customers requesting different types of modes sum up to the total amount of the travel demand; ii) the customers taking shared rides along OD pair $O_k \rightarrow D_k$ is equal to the summation of O_k as the first and second pick up location from all shared-ride trips. The above module is to minimize the total disutility of all customers. Since V_k^m , V_{kk}^m , V_{kk}^m are all exogeneous to *Module II* (as they are independent of Q_k^m , $Q_{k'k}^m$, $Q_{k'k'}^m$), our formulation ensures that each customer chooses the mode with the least disutility. An analogy to this is that the shortest path search problem on a transportation network can be reformulated to a linear program to minimize the total cost of all users of the network; more discussions on this can be found in (*[6](#page-16-5)*).

3.3 Network congestion module (*Module III***)**

Because we only consider the congestion effect of ridesorucing services, *Module III* is based on the ridesourcing layer. There are three types of network flow: i) deadhead miles, i.e., the distance CAVs travel from drop-off locations to the next pick up locations, which apply to both single rides and shared rides; ii) detours, when CAVs are occupied by only the first group of shared ride customers to pick up the second group of customers, which only apply to shared rides; iii) occupied trips, when a CAV travels from the final pick-up location to the destination, which apply to both single rides and shared rides. The following equations show how the three types of flow can be calculated.

(i) Deadhead miles: $\sum_{m=1,3} z_{jk}^{m} t_{j\omega_{k}} + \sum_{m=2,4} \sum_{k'} z_{jkk'}^{m} t_{j\omega_{k}}$ (ii) Detours: $\sum_{m=2,4} \sum_{k' \in K} Q_{kk'}^m t_{O_kO_{kt}}$ (iii) Ocuppied trips: $\sum_{m=1,3} Q_k^m t_{O_k D_k} + \sum_{m=2,4} \sum_{k' \in K} Q_{k'k}^m t_{O_k D_k}$

For each type of flow, we could formulate it as complimentary conditions to ensure that the route choice follows the Wardrop's principle. This is similar to (*[6](#page-16-5)*) and details are omitted here. Now that we have the full formulation of the general equilibrium model, consisting of the three optimization problems from *Module I-III*, one for each of the major players. We can derive the KKT conditions of the three optimization problems and reformulate the general equilibrim model to a **mixed complimentarity**

problem (MCP). We have studied and established the solution existence and unique conditions to such a formulation, using similar methods as reported in (*[6](#page-16-5)*). We omit the details here due to space limitations.

4. NUMERICAL EXPERIMENTS

We solve the general equilibrium model (an MCP) using the *PATH* solver in GAMS. Sensitivity analysis is conducted on a small network and the Sioux Falls network. Due to space limitations, we omit the results of the Sioux Falls network in this paper. When conducting the sensitivity analysis, we first set the baseline values for all parameters; we then either unilaterally changing one parameter at a time or changing two parameters at the same time to see how customer's mode choice and the overall network congestion react to the change of the parameters.

The link properties of the small network (**Figure 3**) are shown in **Table 1**. The demand from node 1-3 to destination node 5 is 40, 40, 40, respectively. **Table 2** lists the basedline parameters. Essentially, parameters for all single ride vehicles are the same, either for whole trips or first mile trips. We set the single-ride parameters for mode 1,3 to the same values, e.g., $\alpha_1^1 = \alpha_1^3$. Similarly, we set the shared-ride parameters for mode m=2,4 to the same values, e.g., $\alpha_1^2 = \alpha_1^4$. These parameters are utility terms. The relative value of a parameter for different modes measures the relative cost of the modes or the relative value of time of customers. The baseline parameter setting in **Table 2** depicts the scenario when the fare of the services follows the relation: single ride > shared ride > transit, and customers inconvenience cost follows the relation: shared ride > single ride. When conducting the sensitivity analysis later in this section, we may change some of the baseline parameters.

TABLE 1 Parameter of the small network

TABLE 2 Baseline parameters

4.1 Results of unilaterally changing one parameter

Table 3 shows the results when we unilaterally change the travel distance based fare rate of single rides or shared rides. Under the current baseline parameter setting (**Table 2**), the demand for serving the whole trip using ridesourcing $(m = 1,2)$ is always zero, indicating that customers prefer the integrated modes that contain both ridesourcing and transit. Compared with Ttransit, RStransit saves more VMT, thus higher demand of RStransit corresponds to lower VMT and vehicle hours traveled (VHT). When we unilaterally increase the distance-based fare rate of single rides $(\alpha_2^1 = \alpha_2^3)$, the cost of requesting Ttransit increases, thus the demand for Ttransit decreases while the demand for RStransit increases. When we unilaterally increase the distance-based fare of shared ride $(\alpha_2^2 = \alpha_2^4)$, the demand of RStransit decreases while the demand of Ttransit increases. This example shows when we increase the cost parameter related to a particular mode (which represents certain cost/disutility of selecting the mode), customer choice of that mode will decrease, and the VMT and VHT of the entire network will also change accordingly.

TABLE 3 Unilaterally change α_2^1 or α_2^2

When we unilaterally change the waiting time cost parameter for single rides (γ_2^1) , modes 1,3,4 have non-zero demand (**Table 4**). When we increase γ_2^1 , the cost of both taxi and Ttransit increases. As a result, the demand of taxi and Ttransit decreases while the demand of RStransit increases.

TABLE 4 Unilaterally change

Table 3-**4** describe the simplest case for the sensitivity analysis. The mode split pattern is not always intuitively predictable, especially when three or more types of modes have non-zero demand simultaneously. **Table 5** summarizes the results when unilaterally change the transfer cost parameter of transit (y_5). When the transfer cost of transit increases, it is intuitive that the demand of RStransit decreases. However, the demand of Ttransit increases despite the increased transit cost. In fact, there is a demand increase in both of the single-ride modes, i.e., taxi and Ttransit. Remember that under equilibrium, a customer always chooses the mode with the lowest disutility. Under the current parameter setting, the fare of single rides is higher than shared rides. The cost of the integrated modes (m=3,4) is the

sum of the cost of the first-mile cost and the transit cost. Therefore, the transit cost takes up higher percentage of the total cost of RStransit compared with Ttransit. When we increase the cost of transit, the disutility of choosing RStransit increases a lot. Since RStransit is no longer a low-cost option for many customers, its demand decreases. In comparison, the transit cost only takes up a small portion of the disutility of choosing Ttransit. Therefore, when we increase the transit transfer cost, the disutility of choosing Ttransit only increases a little bit. Compared with RStransit, Ttranist is still the lower cost option for many customers. Thus more customers choose Ttransit. However, if we further increase the transfer cost of transit, no one chooses the integrated mode (m=3,4), and customers only choose taxi or ridesharing instead for the entire trip.

TABLE 5 Unilaterally change

4.2 Results of changing two parameters at the same time

To better show how different mode choices may impact VMT, we compare the VMT of a certain mode split scenario to the case when all customers drives their private vehicle to work (i.e. all are solo driving). When all customers choose solo driving (no ridesourcing or transit involved), the VMT is equal to 1266.11 vehicle miles based on the UE solution. Thus the VMT change can be calculated as: *VMT change = (VMT of a particular scenario –* 1266.11)/1266.11

VMT change is a better indicator of congestion level compared with the absolute value of VMT. When VMT change is positive, it indicates that there are deadhead miles traveled, which make the congestion level higher than the solo driving scenario. When VMT change is negative, it implies that some customers choose transit or shared rides, so that the network is less congested than the solo driving scenario.

Figure 4 shows the results when we change the time-based disutility rate for shared ride $(\beta_1^2 =$ β_1^4) and the transfer cost of transit (γ_5) at the same time. The demand of single rides is zero. Customers choose between ridesharing amd RStransit. Since the demand of ridesharing and RStransit sum up to 1, we only show the demand pattern of ridesharing (**Figure 4 (a)**). **Figure 4** is divided into 2 regions, with the top part as the high ridesharing demand region, and the bottom part as the high RStransit region. When γ_5 increases, the cost of taking transit increases, thus the demand of RStransit decreases and the demand of ridesharing increases. When β_1^2 increases and the transit parameter $\gamma_5 \approx 0.3$, the cost of ridesharing increases, thus the demand of RStransit increases while the demand of ridesharing decreases. When customers only choose between ridesharing and RStransit, the demand pattern is mainly affected by transit parameter γ_5 . When $\gamma_5 \in [0.25, 0.48]$, the ridesharing parameter also affects the demand pattern. VMT change is consistent with the mode split pattern. VMT gets higher when more customers switch from RStransit to ridesharing.

Figure 4 Sensitivity analysis of changing β_1^2 ($\beta_1^2 = \beta_1^4$) and γ_5

When more than 2 types of modes have non-zero demand, the mode split pattern can be more complex. **Figure 5** shows the results of changing the distance-based fare rate of shared ride $(\alpha_2^2 = \alpha_2^4)$ and the travel distance-based cost of transit (α_3) . The demand of Ttransit is always zero. Customers choose among the other three modes: taxi, ridesharing and RStransit. There are five mode split patterns, marked in **Figure 5 (a)-(b)**. The demand and VMT change of these five mode split patterns are summarized in **Table 6**. Along the boundary of different mode split patterns, the demands of different modes change gradually, shown by the zoom-in version of the boundaries (**Figure 5 (a)-(b)**). Mode split pattern (1) lies in the region where the transit parameter is low, $\alpha_3 < 0.1$. The disutility of taking transit is low, all customers choose RStransit. Mode split patterns ②-⑤ roughly divide the space of transit parameter and ridesharing parameter into four quadrants. Mode split pattern (2) takes place in the right bottom quadrant, where transit parameter $\alpha_3 < 0.64$, shared ride parameter α_2^2 is large. Compared with pattern (1) , transit cost and shared-ride cost are higher in (2) . Thus 33% customers switch from RStransit to taxi. Mode split pattern (3) lies next to (2) , with smaller shared-ride cost parameter. As a result of the decreased shared-ride cost, we see ridesharing demand increases to 67%, taxi demand drop to zero. Compared with (2), mode split (3) also has higher transit cost parameters, thus the transit use decreases to 33%. Mode split ③ has lower VMT compared to ②, since no one request taxi ride in ③. The mode split pattern $\overline{4}$ and $\overline{5}$ can be explained in a similar manner. From this example, we can see that the mode split in the integrated multimodal network has complex patterns. Especially in the era of CAVs, when the operations of vehicles are highly coordinated, it will be even more important to study the customers' mode choice systematically. The VMT change in **Figure 5 (d)** are consistent with the mode split pattern in **Figure 5 (a) – (c)**. The relation between mode split and VMT change is further disccussed in the following section.

Figure 5 Sensitivity analysis of changing α_2^2 $(\alpha_2^2 = \alpha_2^4)$ and α_3

4.3 Model choices versus VMT change

Figure 6 (a) – (d) show the relation between the demand (expressed in terms of the percentage of the total demand) of a particular mode $m \in \{1,2,3,4\}$ and the VMT change. The plots are generated by 1) randomly changing all the model parameters (and thus the mode choices of customers and the resulting VMT of the network); and 2) showing the percentage of the demand *m* and the VMT change. For a given

demand percentage of a mode, for example ridesharing (**Figure 6 (b)**), customers may choose among the other three modes depending on the actual parameter combinations, resulting in different mode choices and different levels of VMT changes. Therefore, the relation between ridesharing demand and VMT change is scattered and forms a triangular region – the variations of mode splits and VMT changes reduce as the demand percentage of ridesharing increases which diminish when the demand percentage is 100% . Similar patterns can be found for the other three modes. Further examinzation of the plots and results reveal that the upper boundary and lower boundary of each plot in **Figure 6** (a) – (d) captures the cases when exactly 2 modes are selected. For example, in **Figure 6 (d)**, along the upper boundary, customers choose between taxi and RStransit, while the demand for ridesharing or Ttransit is zero. Thus, the VMT change starts at the corner case of all customers choosing taxi (VMT change $= 100\%$), ends at the corner case of all customers choosing RStransit (VMT change≈ −80%). Along the lower boundary in **Figure 6 (d)**, customers choose between Ttransit and RStransit while the demand for taxi or ridesharing is zero. The VMT change starts at the corner case of all customers choosing Ttransit (VMT change≈-64%), ends at the coner case of all customers choosing RStransit (VMT change≈ −80%). The data points within the triangular region are cases when 3 or more types of modes are selected.

We can further show the corner cases (i.e. when customers only choose one of the four modes) in **Figure 6(e)**. The plot shows that RStransit save the most VMT: VMT decreases by 80% compared with solo driving when all customers choose RStransit. One the other hand, Ttransit reduces VMT by 64%, ridesharing has similar VMT with solo driving, taxi causes the most congestion with 100% VMT increase due to deadhead miles.

(e) Summary of corner cases

Figure 6 Demand versus VMT change

DISCUSSIONS

In the numerical experiments in section 4, we change the parameters of the four modes, including the fixed fare, time-/distance based fare, time-/distance based disutility, customer waiting time cost, transfer cost, and matching cost. In reality, customers choose one mode over another based on various factors. By setting different parameters, we can better reveal customers' mode choices. It is intuitive that when we increase the parameter (representing certain cost) of a mode, fewer customers will choose that mode. Our results can capture this properly (**Table 3-4**). The proposed model can also capture more complex scenarios due to the interactions of different modes. For example, when we change the parameters of transit, the disutility of both Ttransit and RStransit changes. When we change the parameters of shared rides, the disutility of both ridesharing and RStransit changes. This resembles the interactions of ridesourcing company and transit agencies in practice. When a ridesoucing company changes the price of its shared-ride service, it may affect ridesharing and RStransit differently. Similarly, the disutility due to transit tranfer may change when the transit company make changes to its routes, stops, schedules, and fare. The increased transfer cost of taking transit may affect Ttransit and RStransit differently. Therefore, when there are changes to the parameters of transit and shared rides, how customers exactly react to such changes can be quite complex, which can be captured by the model proposed in this paper as shown in **Tables 3-5** and **Figures 4-5**.

On the other hand, different mode splits (resulted from customers' mode choices) may lead to varied VMT changes (compared with all customers solo driving). Numerical experiments (**Figure 6(e)** in particular) show that higher usage of transit or shared rides reduces the VMT. Shared ride alone without transit (i.e. the ridesharing mode (m=2) in this paper) may still lead to slightly higher VMT compared with solo driving due to the deadhead miles. The results thus clearly illustrate the important role of transit in serving commuters in urban areas. Even we only consider fixed route, fixed schedule mass transit (which are reviewed as less efficient compared with on-demand transit) in this paper, the numerical results show that if all customers choose to use transit, the VMT reduction can be tremendous: 64% if the first mile is served by single rides (for Ttransit) or 80% if the first mile is served by shared-rides (for RStransit), as shown in **Figure 6 (e)**. The numerical results in Section 4 are for the morning commute where travel demands are extremely asymmetric, i.e., all demands are from the residential areas to the worksites. This also explains why the VMT changes are large in general, from 80% reduction (RStransit only) to 100% increase (Taxi only). For more symmetric demand patterns, we expect the VMT changes are milder; see (*[6](#page-16-5)*). However, we believe that the general trend of the changes should be the same.

Our proposed model can thus help provide insights to transit agencies and ridesoucing companies for making sensible policies related to their operations and for the potential collaboration on integrated ridesourcing and transit system. The ability to capture the network-wide congestion effect is also a highlight of the model since the evalution of the congestion level in a multimodal network is important especially CAVs are widely deployed.

CONCLUSIONS

The era of CAV may bring about a significant reduction in car ownership. This paper envisions a morning commuting scenario when there is one ridesourcing platform operating all the vehicles, providing single- or shared- ride services for commuters. CAVs can either send customers to their worksites or send customers to transit stations to take transit to the worksites. This resulted in four modes for a customer to choose: taxi, ridesharing, Transit and RStransit. We modeled the value of time/inconvenience of customers as an overall disutility by choosing a mode, encompassing various factors such as distance, waiting, matching, transfer and the fare of the mode. At equilibrium, the ridesourcing company maximizes its profit, each customer chooses the mode with the lowest disutility, and the vehicular flow is assigned to the network routes based on the Wardrop's first principle, revealing the network congestion. The proposed model captures the behavior and interactions of the ridesoucing platform and the customers and can assess the congestion level of the network. Future studies should test the proposed model on large, real world networks and extend the current model to include other ridesourcing modes such as on-demand transit (i.e. microtransit) and bikeshare services.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Fan, MacCabe, Ban; data collection: Fan; analysis and interpretation of results: Fan, MacCabe, Ban; draft manuscript preparation: Fan, MacCabe, Ban. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- 1. Ke, J., Xiao, F., Yang, H. and Ye, J., 2019. Optimizing Online Matching for Ride-Sourcing Services with Multi-Agent Deep Reinforcement Learning. arXiv preprint arXiv:1902.06228.
- 2. Schaller Consulting, 2018. Estimating Uber and Lyft Ridership in the United States [Accessed: Apr, 2020] <http://schallerconsult.com/rideservices/unsustainable.htm> .
- 3. Shaheen, S., Chan, N., Bansal, A. and Cohen, A., 2015. Shared mobility: definitions, industry developments, and early understanding. University of California Berkeley Transportation Sustainability Research Center, Berkeley.
- 4. Li, W., Cui. Z., Li, Y., Ban, X., 2019. Characterization of ridesplitting based on observed data: A case study of Chendu, China. Transportation Research Part C 100, 330-353.
- 5. Meyhofer, E., 2018. Uber and Toyota Team Up on Self-Driving Cars: Uber Newsroom US. Uber Newsroom. Available at: [Accessed: Jun. 2020] [https://www.uber.com/newsroom/uber-toyota-team](https://www.uber.com/newsroom/uber-toyota-team-self-driving-cars/)[self-driving-cars/](https://www.uber.com/newsroom/uber-toyota-team-self-driving-cars/)
- 6. Ban, X.J., Dessouky, M., Pang, J.S. and Fan, R., 2019. A general equilibrium model for transportation systems with e-hailing services and flow congestion. Transportation Research Part B: Methodological, 129, pp.273-304.
- 7. Schaller Consulting, 2017. Unsustainable? The growth of app-based ride services and traffic. Travel and the Future of New York City [Accessed: Apr. 2020] <http://schallerconsult.com/rideservices/unsustainable.htm> .
- 8. Stein, D.M., 1978. Scheduling dial-a-ride transportation systems. Transportation Science, 12(3), pp.232-249.
- 9. Black, A., 1995. Urban mass transportation planning.
- 10. Malucelli, F., Nonato, M. and Pallottino, S., 1999. Demand adaptive systems: some proposals on flexible transit. In Operational research in industry (pp. 157-182). Palgrave Macmillan, London.
- 11. Feigon, S. and Murphy, C., 2016. Shared mobility and the transformation of public transit (No. Project J-11, Task 21).
- 12. Chen, P.W. and Nie, Y.M., 2017. Connecting e-hailing to mass transit platform: Analysis of relative spatial position. Transportation Research Part C: Emerging Technologies, 77, pp.444-461.
- 13. Ma, T.Y., Rasulkhani, S., Chow, J.Y. and Klein, S., 2019. A dynamic ridesharing dispatch and idle vehicle repositioning strategy with integrated transit transfers. Transportation Research Part E: Logistics and Transportation Review, 128, pp.417-442.
- 14. Pinto, H.K., Hyland, M.F., Mahmassani, H.S. and Verbas, I.Ö., 2019. Joint design of multimodal transit networks and shared autonomous mobility fleets. Transportation Research Part C: Emerging Technologies.
- 15. Yan, X., Levine, J. and Zhao, X., 2019. Integrating ridesourcing services with public transit: An evaluation of traveler responses combining revealed and stated preference data. Transportation Research Part C: Emerging Technologies, 105, pp.683-696.
- 16. Di, X. and Ban, X.J., 2019. A unified equilibrium framework of new shared mobility systems. Transportation Research Part B: Methodological, 129, pp.50-78.