# AN ANALYTICAL MODEL OF MANY-TO-ONE CARPOOL SYSTEM PERFORMANCE UNDER COST-BASED DETOUR LIMITS

**Xin Dong**
Department of Civil and Environmental Engineering
The Pennsylvania State University
217 Sackett Building
University Park, PA 16802
xjd5036@psu.edu

**Hao Liu*, Corresponding Author**
Department of Civil and Environmental Engineering
The Pennsylvania State University
217 Sackett Building
University Park, PA 16802
hfl5376@psu.edu

**Vikash V. Gayah**
Department of Civil and Environmental Engineering
The Pennsylvania State University
231L Sackett Building
University Park, PA 16802
gayah@engr.psu.edu
phone: 814-865-4014

Word Count: 6633 words + 2 table(s) × 250 = 7133 words

Submission Date: April 11, 2024

**ABSTRACT**

Carpooling is an effective strategy to mitigate congestion. By grouping more people into fewer vehicles smartly, total travel distances can be reduced significantly. However, the effectiveness of a carpooling system highly depends on the proportion of interested users that can be successfully matched with each other and the amount of benefits individual users gain from these matches. This paper develops analytical models to estimate these measures for a carpooling system that serves a many-to-one demand pattern, in which travelers share the same basic destination but from different origins. The models provide the expected match rate and average user surplus achieved as a function of the network size, number of users, and travel costs. Different from other works, this model considers the fact that carpool users would be willing to travel under cost-based detour limits instead of fixed detour limits to create a match based on their origins and travel cost considerations. The results are validated using simulation tests and also used to better understand how different parameters might influence the overall performance of a carpooling system.

*Keywords*: Many-to-One, Carpool system, Cost-based detour limits, Flexible-role

## INTRODUCTION

Continually expanding roadway networks and adding new infrastructure are not cost-effective solutions to alleviate congestion and address environmental concerns associated with automobile transportation. It has been shown that these approaches can create more demand for travel which even worsens the problem, a phenomenon known as induced demand (*1*). An alternative solution is to decrease the number of vehicles by increasing vehicle occupancy. With the increase in the popularity of shared mobility in urban transportation networks, carpooling has become a commonly proposed travel mode to help reach this goal. Different from ridesharing which usually includes not only travelers but also a dedicated driver, carpooling is a self-scheduling mode in which travelers agree to share their trips without the need for extra drivers (*2*). Carpooling first appeared in the United States in the 1970s as a solution to the growing oil crisis; however, carpool demand declined after that time and tends to resurface during periods of increased travel costs such as the oil crises in 2005 and financial crises in 2008 (*3*). Thanks to the boom of the internet and the sharing economy, carpooling has been boosted rapidly recently (*4*).

As critical factors considered by stakeholders for the investment in the carpool system, the financial benefits and operational efficiency improvements resulting from carpool systems have drawn abundant research efforts. Numerous studies have explored how various factors – such as demand (*5, 6*), system design (*7, 8*), pricing strategies (*9*), detour or waiting time policies (*10–12*), matching algorithms (*13–15*), etc. – influence the performance of carpool systems.

There exist studies that explored the influencing factors of the carpool system or ride-splitting system's performance using big data: Lehe et al. (*5*) used data from the carpool service (Scoop) to identify signs of increasing returns to scale in a carpool matching system; specifically, as more participants join the carpool system, there is a greater potential for enhanced matching performance and overall benefits. Liu et al. (*6*) utilized ridesharing data from Chicago to demonstrate how an increase in requested shared trips leads to higher matching rates and reduced detour distances. Some agent-based models were also conducted, which simulated the whole carpool process to reproduce the real scene and explore the influence of certain factors for optimization: For example, Masoud et al. (*16*) modeled the carpool system as a bipartite graph and developed a ride-matching algorithm, in which they used a rolling time horizon framework to increase the matching rate. Huang et al. (*17*) proposed a genetic algorithm to solve the multi-objective optimization problem in the carpool system, which helps increase matched passengers and is more efficient. Zhang et al. (*18*) developed a multi-objective optimization model to find the best pricing parameters that protect both the drivers' and the riders' benefits from carpooling by taxi.

Although these data-based and agent-based studies are very detailed and provide precise estimates of performance, they usually require a large amount of input data – such as individual trip data, OD distributions, history trip itinerary, etc. – which usually is challenging to obtain and time-consuming to process and simulate. Analytical models that require less detailed input data have been developed to obtain more general, but less precise, insights about carpooling system's performance. For example, Daganzo et al. (*19*) developed a general analytical model to quantify the cost and performance of dial-a-ride ride-sharing systems, which considers the system status and dynamics. However, the detour distance in this model can be infinite. Other models have been proposed to address this issue. Daganzo et al. (*10*) developed closed-form formulas using system dynamics to quantify the performance of ridesharing with upper limits of detours and proved that taxis produce shorter trip times than the case without limits. Wang et al. (*11*) proposed a model that can predict the matching probability, expected detour, and shared distance for each passenger.

Their model can be used by carpool companies to predict the profit and develop discount policies to attract riders. Ouyang et al. (*12*) developed a many-to-many carpool analytical model with fixed detour limits and waiting time restrictions. The proposed model can predict the performance of the carpool system in terms of Vehicle Kilometers Traveled (VKT), and Personal Kilometers Traveled (PKT) in idealized settings accurately.

These previous analytical studies assumed that carpool users can be potentially matched within fixed detour limits (a fixed maximum detour distance); however, this assumption is questionable in practice. For instance, if some passengers have very short trip distances, and the assigned drivers experience a quite large detour distance, the payment by the rider would not adequately cover the additional costs incurred from the detour by the driver. In such cases, carpooling would create a negative surplus and increase the total travel cost. Therefore, it's plausible to assume that a driver's acceptable detour distance hinges on the monetary benefits from matching, and varies with trip-specific factors like origin, destination, and the rider's willingness to pay.

In light of the above, the objective of this research is to develop an analytical model to assess a carpool system's performance, accounting for individual variations in acceptable detour distance based on travel cost. The model is used to estimate the match rate and financial benefits in a many-to-one demand pattern—emulating morning commutes from various suburban locations to a central business district. Given that only 8.8% of commuters carpool according to the United States Census Bureau (*20*), this paper concentrates on 'small' carpooling systems with few riders/drivers. Further, the accuracy of the proposed analytical models is demonstrated with simulation results. Additionally, this study investigates the impact of factors like demand, payment-cost ratio, driver-rider ratio, and role flexibility on the carpool system's performance.

The remainder of this paper is organized as follows. The subsequent section provides a detailed description of the realistic scenarios in which this model is applied. Following that, the next section delves into the methodology employed to evaluate the performance of the carpool system. The numerical validation of the analytical model is presented in the next section. The following section presents a thorough analysis of the impact of certain factors on the carpool system's performance using models derived from the study. Finally, the key takeaways and concluding remarks are provided.

## SCENARIO

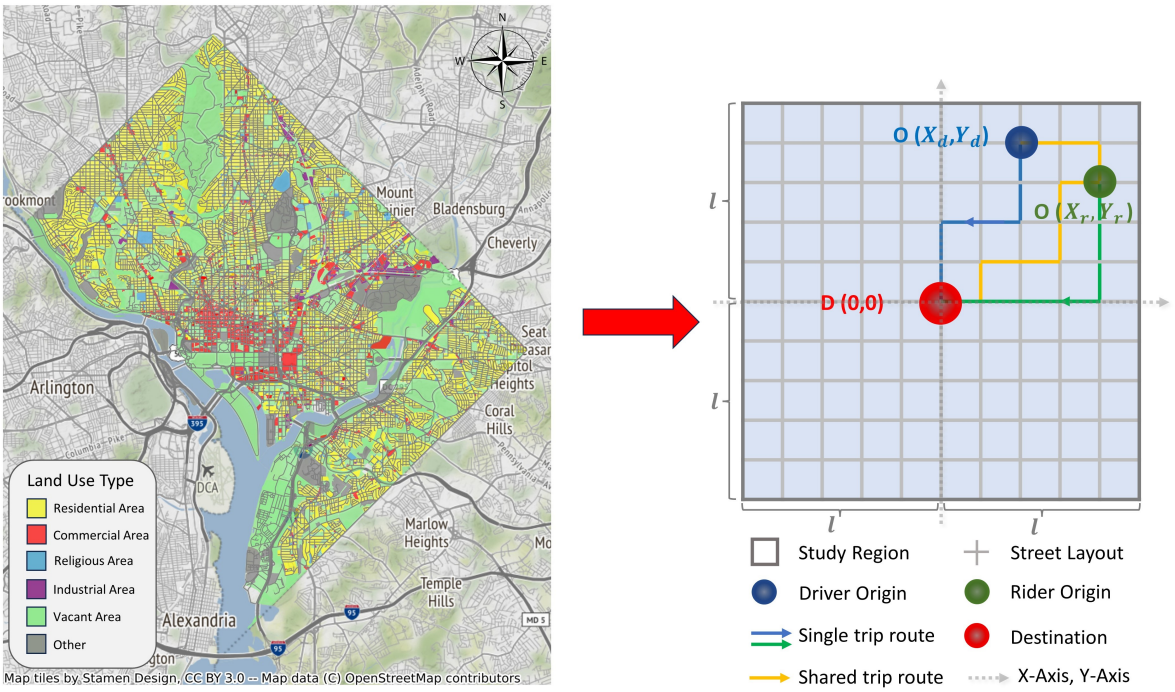
This paper considers a "many-to-one" scenario in which trip origins are homogeneously distributed across a large region but ultimately destined for the same location. We assume that travelers in this system can take one of two roles: a driver or a rider. After the carpool requests are set by users, the system will assign carpool pairs based on whether the shared trip is financially beneficial for both, as will be described later in this section. For simplicity, we assume that each driver can pick up at most one rider during a trip. The rest of this section explains the setting and matching mechanics in more detail.

### Many-to-One Environment

The "many-to-one" environment is common for urban transportation networks; e.g., commuters from all over a city travel towards a dense downtown region during morning peak hours. For simplicity, a city with a dense grid-like street network is considered, and the origins of commuters are assumed to be uniformly distributed across the city. Figure 1(a) presents the comprehensive future land use plan of Washington, D.C. (developed with data from (*21*)), which serves as an

example of the "many-to-one" environment. In this environment, it is expected that the central high-density commercial region (red-colored) attracts commuters from the surrounding residential areas (yellow-colored) during peak hours.

In the analytical model, we considered the environment setting as a $2l\,(km) \times 2l\,(km)$ square region with area $A$, as shown in Figure 1(b). The destinations are assumed to be located at the center of the region with coordinates $(0,0)$, shown by the red dot; the origins of commuters are assumed to be uniformly distributed from the whole area ($A$), examples are shown by the green and blue dots.




(a) Washington D.C. Layout and Land Use (b) An example network of the Analytical Model

**FIGURE 1 Many-to-One Carpooling Environment Setting**

Note that we assume all trips are generated during the same time period (e.g., 15 minutes), and the proposed model can be used to estimate performance metrics within each of these time periods. Therefore, any pair of rider and driver can be matched if they satisfy the criteria defined in the following. The temporal effect on the performance of the carpool system is ignored in this work.

## Trip Routing and Trip Distance

For the network configuration shown in Figure 1(b), the $L_1$ (or Manhattan) metric is employed to calculate the travel distance for all trips. We assume all travelers use the shortest path to complete their trips. Also, we assume if a rider fails to be matched by a driver, the rider will drive alone to the destination.

The blue and green routes in Figure 1(b) are two feasible routes for travelers originating from the blue and green dots if they travel alone. The trip distance for the driver and the rider when

traveling alone can be expressed as $L_d = |X_d| + |Y_d|$ and $L_r = |X_r| + |Y_r|$, respectively.

The yellow route shows a feasible path under a pooling scenario if the driver at $(X_d, Y_d)$ matched with the rider at $(X_r, Y_r)$. The new trip would consist of two parts: 1) the trip from the driver's origin, $O_d$, to the rider's origin, $O_r$, with a length of $L_{dr}$; and, 2) the trip from the rider's origin $O_r$ to their destination $D(0,0)$ with a length of $L_r$. Compared to driving alone, the extra travel distance incurred by the driver, $L_{dt}$, is referred to as detour distance. $L_{dr}$ and $L_{dt}$ can be expressed as:

$$L_{dr} = |X_d - X_r| + |Y_d - Y_r| \tag{1}$$

$$L_{dt} = L_{dr} + L_r - L_d = |X_d - X_r| + |Y_d - Y_r| + (|X_r| + |Y_r|) - (|X_d| + |Y_d|) \tag{2}$$

**Trip Cost**

For simplicity, we estimate all trip costs based on a fixed unit cost per distance, $\alpha(\$/km)$. Further, assume that riders reimburse drivers for a portion of their trips, and the unit cost per distance of this reimbursement, $\beta(\$/km)$, is less than $\alpha$. Based on the unit cost, the following equations are used to describe the various types of expenses involved:

$$C_d^a = \alpha L_d = \alpha(|X_d| + |Y_d|) \tag{3}$$

$$C_d^p = \alpha(L_{dr} + L_r) - \beta L_r = \alpha(|X_d - X_r| + |Y_d - Y_r| + |X_r| + |Y_r|) - \beta(|X_r| + |Y_r|) \tag{4}$$

$$C_r^a = \alpha L_r = \alpha(|X_r| + |Y_r|) \tag{5}$$

$$C_r^p = \beta L_r = \beta(|X_r| + |Y_r|) \tag{6}$$

where $C_d^a$ is the cost a driver incurs without picking up a rider; $C_d^p$ is the cost the driver incurs when picking up a rider, including all detours and reimbursements; $C_r^a$ is the cost the rider incurs if they were to drive themselves; and, $C_r^p$ is the cost the rider incurs in a pooled trip.

**Surplus**

We define the cost savings achieved through carpooling compared to the scenario where individuals travel alone as the 'surplus', which can be used to evaluate the economic benefits produced by the carpool system. The surplus for the driver, $S_d$, and the rider, $S_r$, can be expressed as:

$$\begin{aligned} S_d = C_d^a - C_d^p &= \alpha L_d - [\alpha(L_{dr} + L_r) - \beta L_r] = \beta L_r - \alpha(L_{dr} + L_r - L_d) \\ &= \alpha(|X_d| + |Y_d|) - (|X_d - X_r| + |Y_d - Y_r|)) - (\alpha - \beta)(|X_r| + |Y_r|) \end{aligned} \tag{7}$$

$$S_r = C_r^a - C_r^p = \alpha L_r - \beta L_r = (\alpha - \beta)(|X_r| + |Y_r|) \tag{8}$$

Also, we use $S(X_r, Y_r, X_d, Y_d)$ to refer to the combined surplus of the driver and the rider in a carpool pair (i.e., $S_d + S_r$). And it is dependent on the locations of the driver and the rider, which can be expressed as:

$$S(X_r, Y_r, X_d, Y_d) = S_d + S_r = \alpha(|X_d| + |Y_d| - |X_r - X_d| - |Y_r - Y_d|) \tag{9}$$

**Financially-Feasible Matching Rule**

A financially feasible matching rule is proposed based on the notion of individual rationality in the carpooling system. This rule ensures that carpools would only form (i.e., a driver agrees to pick up a specific rider) when both individuals benefit from the match. Mathematically, this means that carpools only form if $S_r \geq 0$ and $S_d \geq 0$, and this also promises $S(X_r, Y_r, X_d, Y_d) \geq 0$.

Note that engaging in a carpool pair is always financially beneficial for riders under the setting of $\beta < \alpha$. However, it is not always true for drivers. They will only engage in carpooling if the extra cost incurred from the detour can be reimbursed by the rider's payment, which is $S_d \geq 0$.

Compared to a commonly used rule in the literature that assumes drivers are willing to travel with fixed detour limits (same value for everyone) (*12, 22, 23*), denoted by $\Delta$, the cost-

based detour limits (varying value for everyone) in our rule is more reasonable. The fixed detour tolerance assumption requires $L_{dt} \in [0, \Delta]$, which results in $S(X_r, Y_r, X_d, Y_d) \in [\alpha(L_r - \Delta), \alpha L_r]$. Under this scenario, the lower bound of the surplus can be negative if $L_r < \Delta$, which leads to a negative surplus. Thus, the fixed detour tolerance is not always financially feasible.

For this reason, the mechanism we proposed provides a more realistic depiction of carpooling matching, where every participant would find it worthy to join, and promises individual rationality and a self-sustaining carpool system.

## ANALYTICAL MODEL

In this section, we derive analytical models to estimate the performance of the proposed carpool system under various conditions. The measures of performance are the match rate of the system and the system surplus that can be achieved through carpooling arrangements. Two scenarios are considered: 1) a fixed-role scenario in which each traveler can only take a single role (either driver or rider); and, 2) a flexible-role scenario in which travelers can take either role. The models consider the influence of trip demand, as well as the payment-cost ratio, defined as $\gamma = \beta/\alpha$, which represents the relationship between the payment and cost rate. And noted that in our model, we assume $\beta \leq \alpha$, therefore $0 < \gamma \leq 1$.

### Feasible Region for a Rider

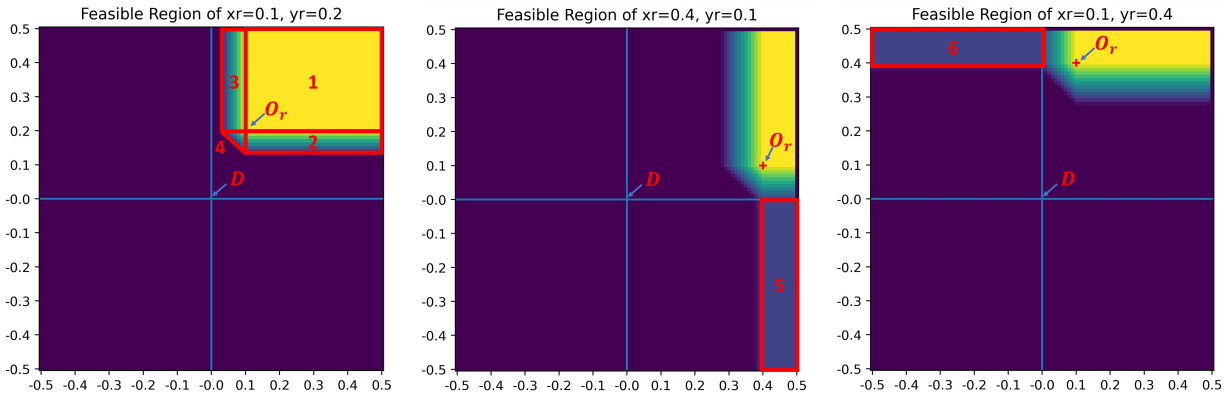
We first define the set of locations in which a driver would be financially motivated to match with a specific rider. We call this set of locations the feasible region for that rider, denote it $\Omega_r(X_r, Y_r)$, and define its area as $A_f(X_r, Y_r)$. Since the network is symmetric and the travelers are assumed to be homogeneously distributed in the network, without loss of generality, we assume the rider is in the first quadrant, i.e., $X_r \geq 0$ and $Y_r \geq 0$. The feasible region for a given rider located at $(X_r, Y_r)$ can be expressed as:

$$\Omega_r(X_r, Y_r) = \{(X_d, Y_d) | (|X_r - X_d| + |Y_r - Y_d|) - (|X_d| + |Y_d|) + (1-\gamma) * (X_r + Y_r) \leq 0\} \quad (10)$$

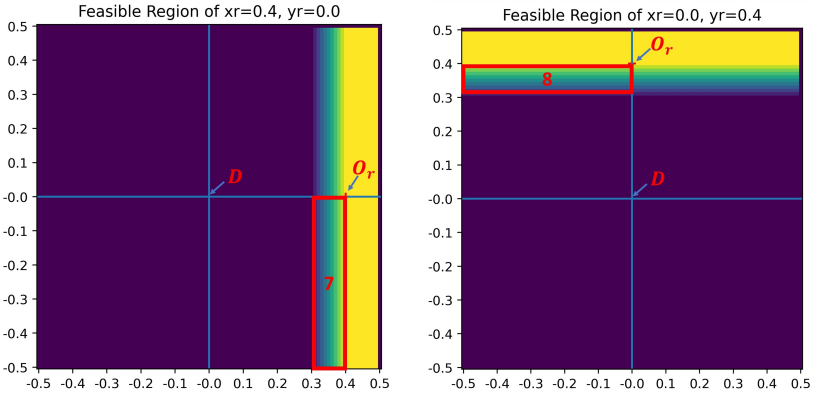
which is derived from Equation (7) by ensuring the non-negativity of $S_d$.

Figure (2) illustrates how this feasible region changes with the rider's location, the example is given with the region area as a unit one. Based on the absolute value sign in Equation (10), the feasible region can be divided into eight sub-regions $\Omega_{r,i}(X_r, Y_r)$, $i = 1, 2, ..., 8$. Note that sub-regions 1~4 in Figure 2(a) exist for every origin of the rider while sub-region 5~8 in Figure 2(b)-2(e) only exist for certain rider origins.

Detailed information for these sub-regions is provided in Table 1: including requirement for the rider's origin, partition of sub-regions expressed with feasible drivers' locations, area and the integration of surplus over each sub-region.




(a) $\Omega_r(0.2, 0.1)$ (b) $\Omega_r(0.4, 0.1)$ (c) $\Omega_r(0.1, 0.4)$




(d) $\Omega_r(0.4, 0)$ (e) $\Omega_r(0, 0.4)$

**FIGURE 2 Feasible Regions ($\Omega_r$) at Different Locations (Sub-regions 1-4 exist at all $(x_r, y_r)s$, while sub-regions 5-8 only exist at certain $(x_r, y_r)s$.)**

**TABLE 1 Feasible Drivers Attributes for a Rider**

| i | Location of $(x_r, y_r)$ | $\Omega_{r,i}$ expressed by $(X_d, Y_d)$ | $A_{f,i}(x_r, y_r)$ | $\iint_{\Omega_{r,i}(x_r,y_r)} S(x_r, y_r, x_d, y_d)$ |
|---|---|---|---|---|
| ① | $x_r \in [0,l]$, $y_r \in [0,l]$ | $X_d \in [x_r,l]$, $Y_d \in [y_r,l]$ | $(l-x_r)(l-y_r)$ | $\alpha(x_r+y_r)(l-x_r)(l-y_r)$ |
| ② | $x_r \in [0,l]$, $y_r \in [0,l]$<br>i. $y_r \geq \frac{\gamma}{2-\gamma}x_r$<br>ii. $y_r \leq \frac{\gamma}{2-\gamma}x_r$ | $X_d \in [x_r,l]$, $Y_d \in [\frac{(2-\gamma)}{2}y_r - \frac{\gamma}{2}x_r, y_r]$ | i. $\frac{\gamma}{2}(l-x_r)(x_r+y_r)$<br>ii. $(l-x_r)y_r$ | i. $\frac{\alpha\gamma(2-\gamma)}{4}(x_r+y_r)^2(l-x_r)$<br>ii. $\alpha(l-x_r)x_r y_r$ |
| ③ | $x_r \in [0,l]$, $y_r \in [0,l]$<br>i. $x_r \geq \frac{\gamma}{2-\gamma}y_r$<br>ii. $x_r \leq \frac{\gamma}{2-\gamma}y_r$ | $X_d \in [\frac{(2-\gamma)}{2}x_r - \frac{\gamma}{2}y_r, x_r]$, $Y_d \in [y_r,l]$ | i. $\frac{\gamma}{2}(l-y_r)(x_r+y_r)$<br>ii. $(l-y_r)x_r$ | i. $\frac{\alpha\gamma(2-\gamma)}{4}(x_r+y_r)^2(l-y_r)$<br>ii. $\alpha(l-y_r)x_r y_r$. |
| ④ | $x_r \in [0,l]$, $y_r \in [0,l]$<br>i. $y_r \geq \frac{2-\gamma}{\gamma}x_r$<br>ii. $\frac{\gamma}{2-\gamma}x_r \leq y_r \leq \frac{2-\gamma}{\gamma}x_r$<br>iii. $y_r \leq \frac{\gamma}{2-\gamma}x_r$ | $X_d \in [0,x_r]$, $Y_d \in [0,y_r]$,<br>$(X_d+Y_d) \geq \frac{2-\gamma}{2}(x_r+y_r)$ | i. $\frac{1}{2}x_r[(\gamma-1)x_r+\gamma y_r]$<br>ii. $\frac{\gamma^2}{8}(x_r+y_r)^2$<br>iii. $\frac{1}{2}y_r[\gamma x_r+(\gamma-1)y_r]$ | i. $\frac{(6\gamma-3\gamma^2-2)}{12}x_r^3 - \frac{(1-\gamma)^2}{2}x_r^2 y_r + \frac{\gamma(2-\gamma)}{4}x_r y_r^2$<br>ii. $\frac{\alpha\gamma(3\gamma-2\gamma^2)}{24A}(x_r+y_r)^3$<br>iii. $\frac{(6\gamma-3\gamma^2-2)}{12}y_r^3 - \frac{(1-\gamma)^2}{2}y_r^2 x_r + \frac{\gamma(2-\gamma)}{4}y_r x_r^2$ |
| ⑤ | $x_r \in [0,l]$, $y_r \in [0, \frac{\gamma}{2-\gamma}x_r]$ | $X_d \in [x_r,l]$, $Y_d \in [-l,0]$ | $(l-x_r)l$ | $\alpha l(l-x_r)(x_r-y_r)$ |
| ⑥ | $x_r \in [0, \frac{\gamma}{2-\gamma}y_r]$, $y_r \in [0,l]$ | $X_d \in [-l,0]$, $Y_d \in [y_r,l]$ | $(l-y_r)l$ | $\alpha l(l-y_r)(y_r-x_r)$ |
| ⑦ | $x_r \in [0,l]$, $y_r \in [0, \frac{\gamma}{2-\gamma}x_r]$ | $X_d \in [0,x_r]$, $Y_d \in [-l,0]$ | $\left[\frac{\gamma}{2}x_r - \frac{(2-\gamma)}{2}y_r\right]l$ | $\alpha l\left[\frac{\gamma(2-\gamma)}{4}(y_r+x_r)^2 - x_r y_r\right]$ |
| ⑧ | $x_r \in [0, \frac{\gamma}{2-\gamma}y_r]$, $y_r \in [0,l]$ | $X_d \in [-l,0]$, $Y_d \in [0,y_r]$ | $\left[\frac{\gamma}{2}y_r - \frac{(2-\gamma)}{2}x_r\right]l$ | $\alpha l\left[\frac{\gamma(2-\gamma)}{4}(y_r+x_r)^2 - x_r y_r\right]$ |

For later derivation of match rate and system surplus, we need the following two terms: the area of the feasible region, expressed as $A_f(x_r,y_r) = \sum_{i=1}^{8} A_{f,i}(x_r,y_r)$; and the expected surplus conditional on $(x_r,y_r)$, $\bar{S}_{x_r y_r X_d Y_d}$, expressed as:

$$\bar{S}_{x_r y_r X_d Y_d} = \frac{1}{A_f(x_r,y_r)} \sum_{i=1}^{8} \iint_{\Omega_{r,i}(x_r,y_r)} S(x_r,y_r,x_d,y_d) dx_d\, dy_d \tag{11}$$

where the integration part of surplus over each sub-region is given in Table 1.

**Fixed-role Carpool System**

This section derives models to estimate the performance of the carpool system when drivers and riders have pre-registered roles. This may exist when each person selects the type of role they wish to perform (either rider or driver) before the matching process.

First, a special case in which only a single rider exists is provided. This is then used to estimate a more general case when $n_r$ riders are present.

*Special Case: Single Rider*

In this case, we assume whether a rider can be matched with a feasible driver is independent of other riders; i.e., at most one rider can locate in the feasible region of a certain driver, if a rider is in the feasible region of the driver, the driver will always agree to be matched with the rider.

*Match Rate*

The probability of a rider being matched when there is 1 rider and $n_d$ drivers is denoted by $p_r(1,n_d)$. This can be calculated as the probability of at least one driver in $\Omega_r(X_r,Y_r)$. Denote the number of drivers that fall within $\Omega_r(X_r,Y_r)$ as $N_\Omega$, then, $p_r(1,n_d) = p(N_\Omega \geq 0 | X_r = x_r, Y_r = y_r)$. The expected value of this probability, denoted as $\bar{p}_r(1,n_d)$, can be expressed as follows:

$$\begin{aligned} \bar{p}_r(1,n_d) &= \bar{p}(N_\Omega \geq 0) = P(X_r = x_r, Y_r = y_r) * \iint_{A_1} p(N_\Omega \geq 0 | X_r = x_r, Y_r = y_r)\, dx_r\, dy_r \\ &= 1 - (\frac{4}{A}) \int_0^{\frac{\sqrt{A}}{2}} \int_0^{\frac{\sqrt{A}}{2}} \left[1 - \frac{A_f(x_r,y_r)}{A}\right]^{n_d} dx_r\, dy_r \end{aligned} \tag{12}$$

where $A_1$ is the first quadrant where all possible $(x_r,y_r)s$ homogeneously distributed; $p(N_\Omega \geq 0 | X_r = x_r, Y_r = y_r)$ can be expressed with $A_f(x_r,y_r)$, provided in Table 1.

It is also crucial to assess the system's match rate, defined as the proportion of travelers that are matched. When there are 1 rider and $n_d$ drivers in the system, we define the system surplus as $\bar{p}_{sys}(1,n_d)$, which can be developed with the rider's expected probability of being matched:

$$\bar{p}_{sys}(1,n_d) = \frac{2\bar{p}_r(1,n_d)}{(1+n_d)} \tag{13}$$

*System Surplus*

The system surplus is defined as the average surplus for all travelers in the system, where both the matched and unmatched agents are considered. The expected value of system surplus when there is 1 rider and $n_d$ drivers is denoted as $\bar{S}_{sys}(1,n_d)$, and expressed in Equation (14). In this scenario, the total surplus is the weighted sum of $\bar{S}_{x_r y_r X_d Y_d}$ given in Equation (11) across all possible riders' locations.

$$\bar{S}_{sys}(1,n_d) = \frac{P(X_r = x_r, Y_r = y_r)}{(1+n_d)} * \iint_{A_1} \bar{S}_{x_r y_r X_d Y_d} * p(N_\Omega \geq 0 | X_r = x_r, Y_r = y_r)\, dx_r\, dy_r$$
$$= \frac{4}{(1+n_d)A} \iint_{A_1} \left[ \frac{1}{A_f(x_r,y_r)} \sum_{i=1}^{8} \iint_{\Omega_{r,i}(x_r,y_r)} S\, dx_d\, dy_d \left(1 - \left(1 - \frac{A_f(x_r,y_r)}{A}\right)^{n_d}\right) \right] dx_r dy_r \quad (14)$$

where S is the abbreviation of $S(x_r, y_r, x_d, y_d)$, which is the surplus of a matched carpool pair, given in Equation (9).

*General Case: Multiple Riders*

The previous section focuses on the case with just one rider in the system. In such a case, if the rider locates in the feasible region of a driver, this driver will always agree to match since there are no other potential matches for the driver. However, when the carpool system gets popular and more riders sign up for the system, it is possible that a driver can receive multiple requests that are financially feasible. In this scenario, the probability of the driver's selection of one specific rider is not 100%, which will greatly change the probability of matching for riders and also the match rate of the system.

Therefore, this section fuses the influence of the driver's selection into formulas of the match rate and system surplus. When there are multiple riders and multiple drivers in the system, we assume the number of each category is $n_r$ and $n_d$, respectively. Also, for simplicity, we assume that the driver has an equal probability of selecting each rider within the feasible region.

*Probability of Selection*

When there are multiple riders that can be matched by the same driver, the probability that the driver selects a specific rider is $p_{select}$. $p_{select}$ is a conditional probability of event $H_1$*: {the driver selecting this specific rider}*, given the event $H_2$*: {a specific rider is feasible for a specific driver}*:

$$p_{select} = \frac{P(H_1 \cap H_2)}{P(H_2)} = \frac{\sum_{n_k=1}^{n_r} \frac{1}{n_k} * \binom{n_r-1}{n_k-1} P^{n_k} (1-P)^{(n_r-n_k)}}{P} \quad (15)$$

where $P$ is an abbreviation of the probability of event $H_2$, calculated with Equation (12) by assigning $n_d = 1$; $n_k$ is the number of riders who locates in the feasible region of the driver, which could range from 1 to $n_r$; $P(H_1 \cap H_2)$ is achieved by aggregating $\frac{1}{n_k}$ weighted by the probability of the occurrence of $n_k$, which is given by the probability of exactly $(n_k - 1)$ successes in the binomial experiment of $B(n_r - 1, P)$.

*Match Rate*

The key distinction between the general case and the special case section is the presence of multiple riders, which leads to competition. $p_{select}$ is applied to estimate the competition and derive the rider's probability of matching and match rate in the general case.

The probability of a rider being matched when there are $n_r$ riders and $n_d$ drivers is expressed as $p_r(n_r, n_d)$, which is the probability of the event *H: {At least 1 driver exists who is in* $\Omega_r(X_r, Y_r)$ *and agrees to match}*. Let $N_{\Omega,yes}$ denote the number of such drivers, then $p_r(n_r, n_d)$ can be expressed as $p(N_{\Omega,yes} \geq 0)$. Therefore, the expected value of this probability, denoted as $\bar{p}_r(n_r, n_d)$, is expressed as:

$$\bar{p}_r(n_r,n_d) = \bar{p}(N_{\Omega,yes} \geq 0) = P(X_r = x_r, Y_r = y_r) * \iint_{A_1} p(N_{\Omega,yes} \geq 0 | X_r = x_r, Y_r = y_r)\, dx_r\, dy_r$$
$$= 1 - \frac{4}{A}\int_0^{\frac{\sqrt{A}}{2}}\int_0^{\frac{\sqrt{A}}{2}} \left[1 - \frac{A_f(x_r,y_r)}{A} * p_{select}\right]^{n_d} dx_r\, dy_r \tag{16}$$

where $p(N_{\Omega,yes} \geq 0 | X_r = x_r, Y_r = y_r)$ is derived with $A_f(x_r,y_r)$ provided in Table 1 and $p_{select}$ provided in Equation (15).

Consequently, the system's match rate when there are $n_r$ riders with $n_d$ drivers, denoted as $\bar{p}_{sys}(n_r,n_d)$, can be expressed as:

$$\bar{p}_{sys}(n_r,n_d) = \frac{2n_r}{(n_r+n_d)}\,\bar{p}_r(n_r,n_d) \tag{17}$$

*System Surplus*

The derivation of system surplus in the general case differs a little from the special case in the incorporation of the probability of selection, $p_{select}$.

The expected value of system surplus when there are $n_r$ riders and $n_d$ drivers, denoted as $\bar{S}_{sys}(n_r,n_d)$, is expressed in Equation (18). In this scenario, the total surplus is the multiplication of the number of riders and the weighted sum of $\bar{S}_{x_r y_r X_d Y_d}$ across all possible riders' locations.

$$\bar{S}_{sys}(n_r,n_d) = \frac{n_r P(X_r = x_r, Y_r = y_r)}{(n_r+n_d)} * \iint_{A_1} \bar{S}_{x_r y_r X_d Y_d} * p(N_{\Omega,yes} \geq 0 | X_r = x_r, Y_r = y_r)\, dx_r\, dy_r$$
$$= \frac{4n_r}{(n_r+n_d)A}\iint_{A_1}\left[\frac{1}{A_f(x_r,y_r)}\sum_{i=1}^{8}\iint_{\Omega_{r,i}(x_r,y_r)} S\, dx_d\, dy_d \left(1 - \left(1 - \frac{A_f(x_r,y_r)}{A} * p_{select}\right)^{n_d}\right)\right] dx_r dy_r \tag{18}$$

**Flexible-role Carpool System**

This section derives measures of performance under the case where travelers do not have pre-assigned roles. In this case, each traveler can be either a driver or rider, depending on if that helps them find a match. This provides more flexibility for the users.

A tree model is utilized to model this flexible system. The structure of this tree is shown in Figure (3). Within the tree, each node has a unique state vector, denoted as $v(v_1,v_2,v_3)$, which encapsulates the number of current waiting agents in the system:

- $v_1$ represents the number of agents in a searching state as riders, and at most 1 agent at a time can be searching for carpool pairs.
- $v_2$ is the number of agents in the driver-only pool who were unsuccessful in finding matches as riders but are waiting for potential matches as drivers.
- $v_3$ presents the number of agents in the flexible-role pool, who has not been assigned a role yet and are available for matching with either role.

Let $N$ be the total number of travelers in the system. The state vector at the root node is $v(1,0,N-1)$, which means from the beginning of the matching process, a traveler is randomly selected to be a rider and searches for a match, regarding the rest $N-1$ flexible agents as drivers. If there is any agent presented in $v_2$, there will be 3 directed edges emanating from this parent node since the matched driver can originate from both $v_3$ and $v_2$; if not, there will be 2 directed edges from it. Then if this searching rider is successfully matched, denoted by the right branches from

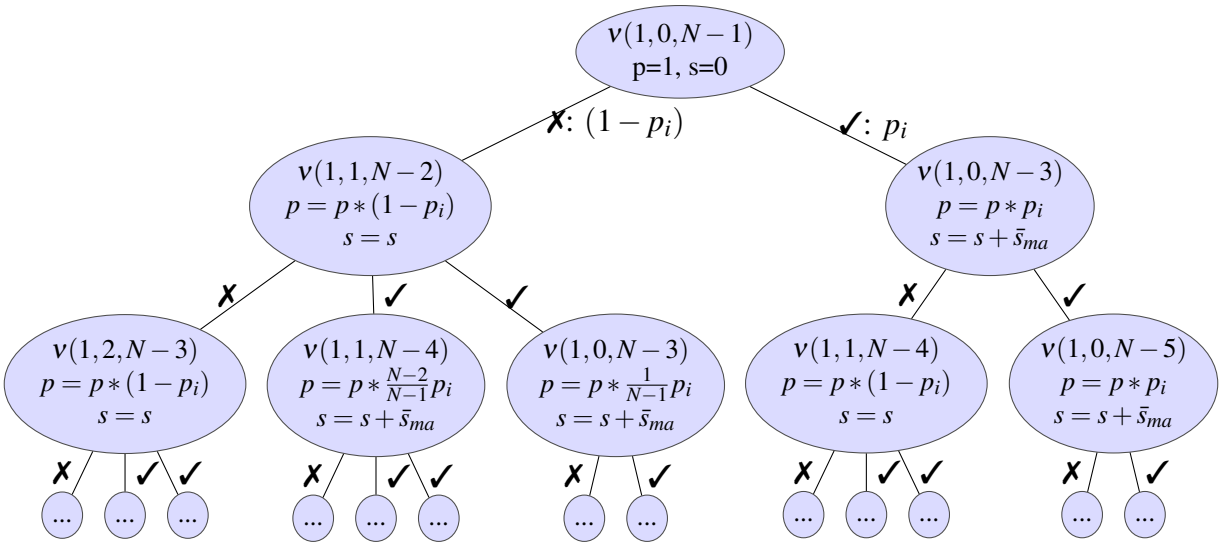



**FIGURE 3 Tree Structure for Flexible-role Carpool Matching System**

the parent node, both the rider and the corresponding driver (selected from either $v_2$ or $v_3$) are removed from the system; however, if a match is not found, denoted the left branch, the searching rider is placed in the driver-only pool, $v_2$, to await potential matches, and one agent out of $v_3$ will be moved to $v_1$ to search. Repeat this process until the system is empty or there are only drivers left, which implies that at the leaf node, $v_1$ and $v_3$ are always 0. For the latter scenarios, all drivers will leave the system by driving alone.

Therefore, the height of this tree is equal to the number of all agents ($N$) within the carpool system, representing the length of the path from the root node to the deepest leaf node. The depth of the leaf tree, $h_i$, varies within $[\lfloor \frac{N}{2} \rfloor, N]$. The maximum depth of the path is $N$ when no one in the system is matched, and the minimum depth is $\lfloor \frac{N}{2} \rfloor$ when the maximum number of travelers is matched.

In the tree, the likelihood shown beside the edge is the conditional probability from its parent node; and $p_i$ is the probability of the $i^{th}$ served agent getting matched with the appearance of $n_i$ drivers, which is the abbreviation of $\bar{p}_r(1,n_i)$. In each node, the probability value, $p$, denotes the likelihood of this current state (following the path from the root node to this node); $s$ denotes the total surplus out of the whole system at this current state. Both $p$ and $s$ are derived iteratively, and updated at each layer of the tree. $\bar{s}_{ma}$ is the expected surplus that a matched pair can achieve, which is given by:

$$\bar{s}_{ma} = \frac{4}{A} \iint_{A_1} \frac{1}{A_f(x_r,y_r)} \iint_{\Omega_r(x_r,y_r)} S(x_r,y_r,x_d,y_d)\, dx_d\, dy_d dx_r dy_r \tag{19}$$

*Match Rate*

With the information from the leaf nodes of the tree model, the system's match rate with $N$ flexible-role agents in the system could be derived, which is denoted as $\bar{p}_{sys}(N)$. This is calculated by dividing the weighted average number of matched agents (weighted by the probability at the leaf node) by the number of agents in the system:

$$\bar{p}_{sys}\ (\mathrm{N}) = \frac{1}{N} \overbrace{\sum_{n_m=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{\sum \delta_i = n_m}}^{\sum \pi} \left\{ \overbrace{(2n_m)}^{n_{leaf}} \overbrace{\prod_{i=1}^{h_i} [\delta_i * \bar{p}_r(1,n_i) + (1-\delta_i) * (1-\bar{p}_r(1,n_i))]}^{pleaf} \right\}$$

$$\delta_i = \begin{cases} 1, & i^{th} \text{ agent matched} \\ 0, & n_i = 0 \text{ or } i^{\text{th}} \text{ agent not matched} \end{cases} \tag{20}$$

$$n_i = \begin{cases} N, & i = 1 \\ min(0, n_{i-1}), & i > 1 \text{ and if } (i-1)^{\text{th}} \text{ agent not matched} \\ min(0, n_{i-1} - 2), & i > 1 \text{ and if } (i-1)^{\text{th}} \text{ agent matched} \end{cases}$$

where $\sum \pi$ controls the sum along all paths by traversing $n_m$, the number of matched pairs in each path, in the range of $[1, \frac{N}{2}]$. $p_{leaf}$ denotes the probability at the leaf node of each path, given by the continuous multiplication of the probability besides the edge along this path, with either $\bar{p}_r(1,n_i)$ (matched) or $1-\bar{p}_r(1,n_i)$ (unmatched). $n_{leaf}$ is the number of matched agents at the leaf node of each path. Other variables are explained as follows:

$\delta_i$ : An indicator whether the $i^{th}$ agent gets matched, and $\sum_{i=1}^{h_i} \delta_i = n_m$.

$n_i$ : The number of potential 'drivers' in the system, which is also the summation of $v_2$ and $v_3$.

$n_m$ : The number of matched pairs along each path, which can not exceed $\lfloor \frac{N}{2} \rfloor$

$h_i$ : The depth of each path, varies within$[\lfloor \frac{N}{2} \rfloor, N]$

$\bar{p}_r(1,n_i)$ : The probability of a rider being matched with the appearance of $n_i$ drivers.

*System Surplus*

The expected system surplus for a flexible-role system with $N$ agents can also be derived with the information from the leaf nodes of the tree model, denoted as $\bar{S}_{sys}(N)$. This value is given by dividing the weighted average of the total surplus (weighted by the probability at the leaf node) by the number of agents in the system:

$$\bar{S}_{sys}(N) = \frac{1}{N} \overbrace{\sum_{n_m=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{\sum \delta_i = n_m}}^{\sum \pi} \left\{ \overbrace{\prod_{i=1}^{h_i} [\delta_i * \bar{p}_r(1,n_i) + (1-\delta_i) * (1-\bar{p}_r(1,n_i))]}^{p_{leaf}} * \overbrace{(\sum_{i=1}^{h_i} \delta_i * \bar{s}_{ma})}^{s_{leaf}} \right\} \tag{21}$$

where $s_{leaf}$ is the total surplus achieved at the leaf node of each path, given by the sum of $\bar{s}_{ma}$, the average surplus of all matched carpool pairs, along the path; $\bar{s}_{ma}$ is calculated with Equation (19).

## NUMERICAL VALIDATION

The validity of the analytical models is tested using agent-based simulations, the parameters for which are outlined in Table 2. Two scenarios are explored:

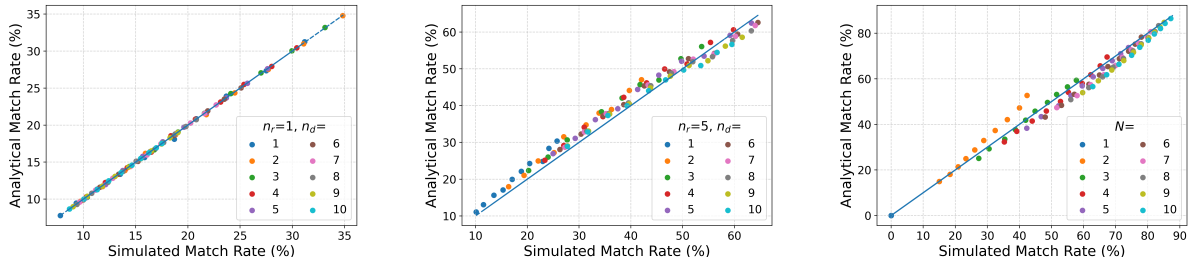
1) For the fixed-role scenario, riders and drivers are randomly generated within the system. A rider is randomly selected to seek carpool pairs among available drivers. If there are multiple feasible drivers, one is randomly selected for matching. If matched, both are removed from the system. Otherwise, unmatched riders exit the system to drive alone, while unmatched drivers remain available for further pairing. Repeat the process until no drivers or riders are left. The match rate is calculated as the number of matched travelers over the number of all. The system surplus is calculated by dividing the sum of all the surplus by the number of agents.

2) For the flexible-role scenario, all travelers are generated at random locations. First, an agent is randomly selected and treated as a rider. The rider searches for a match with other agents in the system, who are seen as drivers by the rider. If there are multiple feasible agents for the searcher, one is randomly selected for matching. If matched, both agents are removed from the system. Unmatched agents stay in the system with a driver-only role, awaiting future potential matches. The process continues until only driver-only agents remain or all possible agents are matched. The remaining agents are assumed to drive alone. Match rates and system surplus are calculated similarly to the fixed-role scenario.

**TABLE 2 Simulation Settings**

| Notation | Description | Value |
|---|---|---|
| A | Network size, the area of the square region ($km^2$) | 100 |
| $l$ | Half the side length of the region ($km$) | 5 |
| $\lambda$ | Trip generating rate of region (per unit time & area) | [0.01,0.2] |
| $n_d$ | Number of drivers in the system | [1,10] |
| $n_r$ | Number of riders in the system | [1,10] |
| $n_f$ | Number of flexible-role agents in the system | [1,20] |
| $\alpha$ | Unit cost for driving alone (\$/km) | 1 |
| $\beta$ | Unit payment being carried (\$/km) | [0.1,1] |
| $\gamma$ | payment-cost ratio, derived by $(\beta/\alpha)$ | [0.1,1] |
| Iteration | Number of simulations with different random seeds | 50000 |

The validation results are illustrated in Figure (4). The three columns from left to right present the fixed-role scenario with 1 rider and 1~10 drivers, the fixed-role scenario with 5 riders and 1~10 drivers, and the flexible-role scenario with 1~10 agents. The horizontal axis depicts the simulated results, and the vertical axis depicts the analytical results. The top row displays the match rate, while the bottom row showcases the system surplus. The different colors in the plot correspond to varying numbers of drivers (travelers) in the fixed-role scenarios (flexible-role scenario). Within each color category, multiple dots are presented, representing different payment-cost ratios.




(a) Match Rate (Fixed-role, $n_r = 1$) (b) Match Rate (Fixed-role, $n_r = 5$) (c) Match Rate (Flexible-role)

(d) Surplus (Fixed-role, $n_r = 1$) (e) Surplus (Fixed-role, $n_r = 5$) (f) Surplus (Flexible-role)

**FIGURE 4 Numerical validation for metrics in the Analytical Model**

Figure 4 shows that the analytical results match well with the simulated results overall. However, in the fixed-role general scenario, the analytical results slightly overestimate the match rate and surplus. A possible reason is the insufficient consideration of the dependence of matching between riders. Remember we use the feasible region, $\Omega$, to capture the probability of a rider being matched. The $1^{st}$ rider's not matching with a driver indicates the driver locates at the complement of the $1^{st}$ rider's feasible region, $\Omega_1'$. Conditionally, whether the $2^{nd}$ rider can be matched with this driver should use the area of $\Omega_1' \cap \Omega_2$, which is a subset of $\Omega_2$. However, in our model, we still use $\Omega$ to capture the probability of the consequent rider being matched with this driver regardless of previous information. Therefore, our analytical model overestimates the system performance a bit in this scenario.

Another observation under the flexible-role scenario, from Figure 4(c), is that when there are more than two agents, the model tends to underestimate the match rate a bit. This is also caused by the insufficient consideration of the dependence of matching between riders. When an agent is not able to be matched as a rider, it is highly possible this agent locates further from the destination (It is obvious in Figure 2 that the further the agent locates, the smaller the feasible region as a rider). Thus this agent actually has a higher potential of being matched as a driver with other agents. But in the tree model, the probability of matching is not captured from this individual level, leading to an underestimation of the match rate. However, the 2 flexible agents case does not follow this trend and the model predicts a higher match rate and surplus. This is because when an agent can not be matched as a rider, it is further from the destination, however, different from multiple agents case, there are not any other agent waiting to be matched with him. The actual probability of being matched as a driver is not greatly lifted up compared with when there are more agents, therefore

the analytical model overestimates.

Overall, the differences between the analytical and simulated results have an acceptable range: For the fixed-role general case (example of 5 riders and 1~10 drivers), on average the absolute difference of match rate observed is 1.4%, and the largest absolute difference is around 5%. The average absolute difference of the system surplus is 0.04$. The average relative difference observed is 8%, and the largest relative difference observed is 19%. For the flexible-role case (example of 1~10 agents), on average the absolute difference of match rate observed is 2.6%, and the largest absolute difference observed is 10%, where the analytical model is underestimating the match rate compared with the simulation. The average absolute difference of system surplus observed is 0.05$, the average relative error observed is 5.5%, and the largest relative error observed is around 24%, which appears when there are only 2 agents.

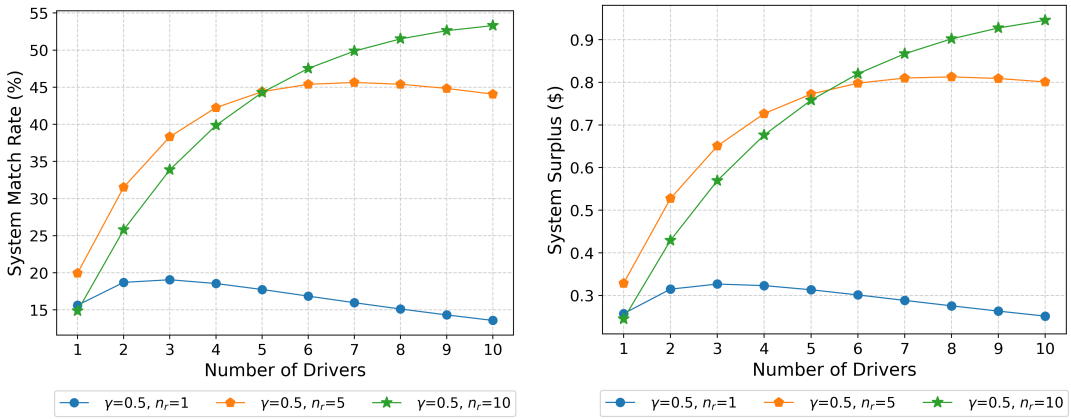
## RESULTS OF ANALYTICAL MODELS

In this section, we examine the impact of various factors on the carpool system's performance with both fixed-role and flexible-role scenarios. Specifically, we focus on understanding how demand, driver-rider ratio, and payment-cost ratio contribute to the overall effectiveness of the system. Additionally, we compare the outcomes between the two distinct scenarios.

### Influence of Demand

Figure 5(a) and Figure 5(b) show the patterns of match rate and system surplus against the number of agents ($n_d$ and $n_r$, respectively). The payment-cost ratio is set to 0.5, while the demand varies.

In both Figure 5(a) and Figure 5(b), the curves generally show a unimodal pattern. Both match rate and system surplus increase with the number of drivers and reach a maximum value when the number of drivers is slightly greater than the number of riders, then start to decrease. The increasing trend occurs when the supply of one side ($n_r$ or $n_d$) is insufficient to meet the demand of the other side ($n_d$ or $n_r$), indicating an imbalance in the system. Conversely, the decreasing trend is observed when the supply exceeds the demand, resulting in an oversupply condition. As the increase and the decrease are caused by the relative relationship between the demand of drivers and riders, it will be elaborated on in the next section.




(a) Match Rate (b) System Surplus

**FIGURE 5 System Performance With the Changes in Demand**

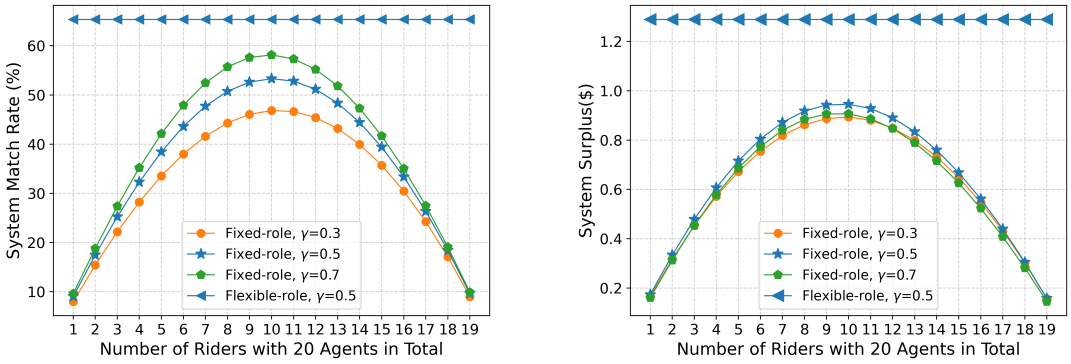
**Influence of Driver-Rider Ratio**

In this section, to explore the influence of the driver-rider ratio on system performance, we fix the total demand of travelers in the system, while changing the driver-rider ratio. In both Figure 6(a) and Figure 6(b), the total number of travelers is set to 20.

Figure 6(a) shows the system performance of the match rate. Along the horizontal axis, the number of riders increases, while drivers decrease in the fixed-role scenario, indicating a decrease in the driver-rider ratio. Three fixed-role scenarios with different payment-cost ratios and one flexible-role scenario are shown in the figure. Notably, the match rate achieved by the flexible-role system (blue horizontal line) surpasses the corresponding fixed-role system with $\gamma = 0.5$ (blue curve). This is due to the flexibility of the agents' roles, which would be elaborated in the last subsection.

In addition, Figure 6(a) reveals that the highest match rate for the fixed-role scenarios is achieved when the driver-rider ratio reaches one under our tested scenarios, indicating the attainment of an equilibrium point. At this ratio, the system is effectively balanced, maximizing the probability of successful matches between drivers and riders. Further, the curves with different colors reveal that changes in the driver-rider ratio have a more pronounced impact when the $\gamma$ is higher. As an example, when $\gamma = 0.7$, the change of the match rate with a unit change if the number of riders is large than that when $\gamma = 0.5$. This can be attributed to the fact that, with a higher payment-cost ratio, the payment to the driver becomes less restrictive in terms of covering drivers' costs. Thus, changing the driver-rider ratio would quickly change the system equilibrium.

Similarly, Figure 6(b) shows the system surplus variations under different scenarios. The system surplus achieved by the flexible-role system surpasses that of the fixed-role system because of the flexibility of the agents' roles. Further, examining each curve within the fixed-role system, the results suggest that the highest system surplus occurs when the driver-rider ratio reaches one under our tested scenarios as well .



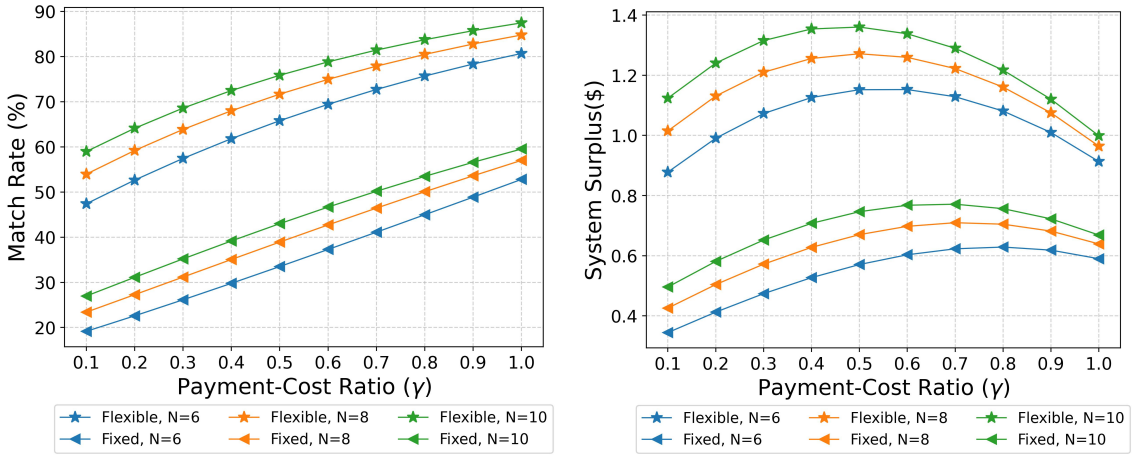

(a) Match Rate (b) Surplus

**FIGURE 6 System Performance With the Changes in Driver-Rider Ratio**

**Influence of Payment-Cost Ratio**

Figure 7(a) and Figure 7(b) depict the relationship between the payment-cost ratio ($\gamma$) and the match rate or system surplus for both the fixed-role and flexible-role systems with different number of agents within the system.

In Figure 7(a), it is obvious that the match rate increases as the payment-cost ratio increases under all scenarios, and the match rate is higher in the flexible role system than that with fixed roles. However, the highest system surplus is not necessarily achieved at the largest payment-cost ratio, observed from Figure 7(b). The system surplus initially rises and starts to decline from a certain point. This phenomenon is a result of the trade-off between the match rate and the detour distance. Increasing the payment-cost ratio leads to an expansion of the feasible region and thus an increase in the probability of matching. However, a larger feasible region also implies a greater distance, on average, between the rider's origin and the driver's, which results in a longer average detour distance. As surplus reduces with the additional distance the driver has to travel, a larger payment-cost ratio leads to a smaller system surplus. Therefore, the trade-off between the match rate and the detour distance explains why the system surplus initially increases but eventually declines as the payment-cost ratio increases.




(a) Match Rate

(b) System Surplus

**FIGURE 7 System Performance With the Changes in Payment-Cost Ratio**

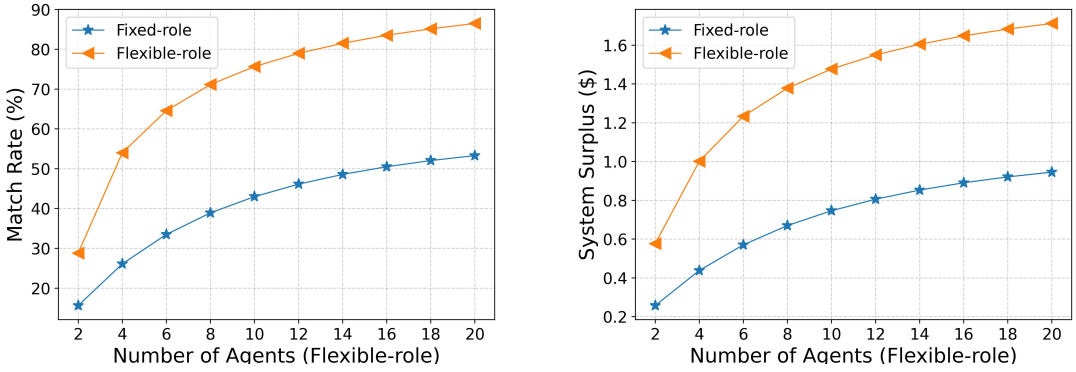
**Comparison of Fixed-role & Flexible-role System's Performance**

Since the above results suggest the driver-rider ratio of one resulted in the highest system performance in the fixed-role system. Therefore, here we use this ratio setting to evaluate and compare the performance of the fixed-role system with that of the flexible-role system.

In Figure 8(a) and Figure 8(b), the driver-rider ratio is set to one, and the payment-cost ratio is fixed as 0.5. The total number of agents ranges from 2 to 20 in both the fixed-role system and the flexible-role system. As can be seen from both figures, the flexible-role system always has better performance than the fixed-role system with the same number of agents in the system. This occurs because of the flexibility offered by travelers being willing to take both roles. In certain scenarios, there may be agents located far from the destination who have a low likelihood of being picked up by other agents but are more likely to function as drivers and pick up riders along their route to the destination. In such cases, the introduction of flexible-role selection would help such agents to match and consequently improve the system's performance.

Observing the slope of both curves in Figure 8(a) and Figure 8(b), the flexible-role system demonstrates a more substantial and rapid increase in the match rate and system surplus as the

demand intensifies from a very small value to a relatively larger value. On the other hand, the curve of the fixed-role system exhibits a comparatively gentler slope, indicating a relatively slower response to increased demand due to the predefined roles and constraints within the system.




(a) Match Rate (b) System Surplus

**FIGURE 8 System Performance of Fixed-role and Flexible-role System**

## DISCUSSION AND CONCLUSION

This study investigates the performance of a carpooling system with cost-based detour limits in a grid-networked squared city with homogeneous many-to-one demand. Analytical models are provided to predict the expected match rate between travelers and the average surplus (benefits) out of the system under two scenarios: one with the agents having fixed-roles (either as drivers or as riders), and the other with flexible-roles (both roles). The results of the analytical models are compared to values obtained from simulations of the proposed carpool system.

Overall, the results from the proposed models match well with the simulated results under various conditions. The models are then used to investigate the impact of certain factors, such as the payment-cost ratio, driver-rider ratio, number of travelers, and the role of flexibility on the match rate and system surplus. The results reveal the benefits of flexible-roles are significant as the fixed-roles greatly reduce the opportunity to create financially feasible matches. This demonstrates the importance of ensuring that travelers are encouraged to seek both roles or motivated (via subsidies or other schemes) to balance between preferred roles. In the fixed-role cases, match rates and surpluses are generally maximized when the number of riders and drivers are equal under our tested scenarios. In conclusion, while the proposed analytical model has limitations and assumptions, it offers valuable insights and trends that can assist stakeholders in assessing the benefits and potential returns on investment in a carpool system.

Future refinement could focus on improving the assumptions made in the model which may not be satisfied in practice, including the trip cost calculation, distribution of the demand, and network layouts, etc. For instance, improving the trip cost's consideration and the financially feasible rule, which currently relies solely on fuel costs. However, it may not accurately reflect the additional cost of time resulted from detouring. To address this limitation, future work could incorporate the cost of time and schedule delays into the trip cost. Another aspect to explore is expanding the many-to-one system to a many-to-many system, with multiple origins and destinations.

This would provide planners and stakeholders with a more comprehensive tool with versatility and applicability.

Furthermore, the model assumes a square city layout with homogeneous demand, which may not accurately reflect the complexities of real-world transportation networks and diverse demand patterns. By incorporating actual network layouts and heterogeneous demand distributions, the model can be enhanced to provide more realistic and robust insights.

## ACKNOWLEDGEMENTS

This research was supported by NSF Grants CMMI-2052337.

## AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: X. Dong, H. Liu, V. Gayah; simulation: X. Dong; analysis and interpretation of results: X. Dong, H. Liu, V. Gayah; draft manuscript preparation: X. Dong, H. Liu, V. Gayah; All authors reviewed the results and approved the final version of the manuscript.

## REFERENCES


1. Cervero, R., Road expansion, urban growth, and induced travel: A path analysis. *Journal of the American Planning Association*, Vol. 69, No. 2, 2003, pp. 145–163.
2. Handke, V. and H. Jonuschat, *Flexible ridesharing: new opportunities and service concepts for sustainable mobility*. Springer Science & Business Media, 2012.
3. Olsson, L. E., R. Maier, and M. Friman, Why do they ride with others? Meta-analysis of factors influencing travelers to carpool. *Sustainability*, Vol. 11, No. 8, 2019, p. 2414.
4. Furuhata, M., M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang, and S. Koenig, Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B: Methodological*, Vol. 57, 2013, pp. 28–46.
5. Lehe, L., V. V. Gayah, and A. Pandey, Increasing Returns to Scale in Carpool Matching: Evidence from Scoop. *Transport findings*, 2021.
6. Liu, H., S. Devunuri, L. Lehe, and V. V. Gayah, Scale effects in ridesplitting: A case study of the City of Chicago. *Transportation Research Part A: Policy and Practice*, Vol. 173, 2023, p. 103690.
7. Regue, R., N. Masoud, and W. Recker, Car2work: Shared mobility concept to connect commuters with workplaces. *Transportation Research Record*, Vol. 2542, No. 1, 2016, pp. 102–110.
8. Zhang, D., T. He, Y. Liu, S. Lin, and J. A. Stankovic, A carpooling recommendation system for taxicab services. *IEEE Transactions on Emerging Topics in Computing*, Vol. 2, No. 3, 2014, pp. 254–266.
9. Zhang, D., Y. Li, F. Zhang, M. Lu, Y. Liu, and T. He, coRide: Carpool service with a win-win fare model for large-scale taxicab networks. In *Proceedings of the 11th ACM conference on embedded networked sensor systems*, 2013, pp. 1–14.
10. Daganzo, C. F., Y. Ouyang, and H. Yang, Analysis of ride-sharing with service time and detour guarantees. *Transportation Research Part B: Methodological*, Vol. 140, 2020, pp. 130–150.
11. Wang, J., X. Wang, S. Yang, H. Yang, X. Zhang, and Z. Gao, Predicting the matching probability and the expected ride/shared distance for each dynamic ridepooling order: A mathematical modeling approach. *Transportation Research Part B: Methodological*, Vol. 154, 2021, pp. 125–146.
12. Ouyang, Y., H. Yang, and C. F. Daganzo, Performance of reservation-based carpooling services under detour and waiting time restrictions. *Transportation Research Part B: Methodological*, Vol. 150, 2021, pp. 370–385.
13. Xia, J., K. M. Curtin, W. Li, and Y. Zhao, A new model for a carpool matching service. *PloS one*, Vol. 10, No. 6, 2015, p. e0129257.
14. Masoud, N. and R. Jayakrishnan, A decomposition algorithm to solve the multi-hop peer-to-peer ride-matching problem. *Transportation Research Part B: Methodological*, Vol. 99, 2017, pp. 1–29.
15. Tafreshian, A. and N. Masoud, Trip-based graph partitioning in dynamic ridesharing. *Transportation Research Part C: Emerging Technologies*, Vol. 114, 2020, pp. 532–553.
16. Masoud, N. and R. Jayakrishnan, A real-time algorithm to solve the peer-to-peer ride-matching problem in a flexible ridesharing system. *Transportation Research Part B: Methodological*, Vol. 106, 2017, pp. 218–236.

17. Huang, S.-C., M.-K. Jiau, and C.-H. Lin, A genetic-algorithm-based approach to solve carpool service problems in cloud computing. *IEEE Transactions on intelligent transportation systems*, Vol. 16, No. 1, 2014, pp. 352–364.
18. Zhang, W., R. He, Y. Chen, M. Gao, and C. Ma, Research on taxi pricing model and optimization for carpooling detour problem. *Journal of Advanced Transportation*, Vol. 2019, 2019.
19. Daganzo, C. F. and Y. Ouyang, A general model of demand-responsive transportation services: From taxi to ridesharing to dial-a-ride. *Transportation Research Part B: Methodological*, Vol. 126, 2019, pp. 213–224.
20. Bureau, U., American community survey (acs). *The United States Census Bureau nd https://www.census.gov/programs-surveys/acs (accessed Jul. 25, 2022)*, 2022.
21. OpenDataDC, Single Member District in D.C. *https://opendata.dc.gov/datasets/single-member-district-from-2023*, 2023.
22. Lin, Y., W. Li, F. Qiu, and H. Xu, Research on optimization of vehicle routing problem for ride-sharing taxi. *Procedia-Social and Behavioral Sciences*, Vol. 43, 2012, pp. 494–502.
23. Zhang, W., R. He, C. Ma, and M. Gao, Research on taxi driver strategy game evolution with carpooling detour. *Journal of Advanced Transportation*, Vol. 2018, 2018, pp. 1–8.