THEMA

théorie économique, modélisation et applications

THEMA Working Paper n°2021-17
CY Cergy Paris Université, France

Social optimality and stability of matchings in peer-to-peer ridesharing

Paolo Delle Site, André de Palma, Samarth Ghoslya

August, 2021

CERGY PARIS UNIVERSITÉ

cnrs

Social optimality and stability of matchings in peer-to-peer ridesharing

Paolo Delle Site°, André de Palma°°, Samarth Ghoslya°°°

° Niccolò Cusano University, Rome, Italy
°° CY Cergy Paris Université, France
°°° Sapienza University, Rome, Italy

Abstract

Peer-to-peer ridesharing, where drivers are also travellers, can alleviate congestion and emissions that plague cities by increasing vehicle occupancy. We propose a socially optimal ridesharing scheme, where a social planner matches passengers and drivers in a way that minimizes travel costs (travel time and fuel) plus environmental costs. The contribution helps in computing the socially optimal ridesharing schemes for networks of any topology within a static framework of route choice with exogenously fixed travel times. A linear programming problem is formulated to compute the optimal matchings. Existence, integrality and uniqueness properties are investigated. The social planner receives a payment from passengers and rewards drivers for the higher costs they bear. Passengers and drivers never incur a loss because travelling alone remains always an option, but matchings may need to be subsidised. The socially optimal matching solution without environmental costs is proved to satisfy the stability property according to which no pair of passenger and driver prefers each other to any of the current partners. In the Sioux Falls network, when 20% of individuals are willing to rideshare, with 80% of passengers travelling by car and 20% by public transport, 17.37% optimally do so, resulting in a 7.05% decrease in CO_2 emissions on the all-travel-alone scenario.

Keywords: environment, matching stability, optimization, ridesharing, socially optimal matching
JEL classification: C78, R40, R48

1. Introduction

Ridesharing, also referred to as carpooling, increases vehicle occupancy by allowing more individuals to travel together while sharing their travel costs. Ridesharing is expanding its market penetration as innovative mobility service in cities and between cities and became more prevalent worldwide.

Ridesharing services can be classified as follows. The first type is informal ridesharing. The second type is community ridesharing. This ridesharing involves a third party providing the matching functionality (e.g. Zimride). Examples of community ridesharing are home-to-work travel by colleagues of the same or neighbouring company, or home-to-event travel by event participants (e.g. conference participants). The third type is peer-to-peer ridesharing, where drivers are also travellers (e.g.

Blablacar). The provision of the service by drivers is not for profit aims. The platform may charge a commission fee, which adds to the transfer payment between passenger and driver. The fourth type is ridesharing where drivers are not travellers and provide the service for profit aims. Uber is an example of a Transportation Network Company (TNC) providing matching between such drivers and passengers (Campbell, 2018). A new job market has been created. Opposition from taxi corporations, especially when taxis pay for an expensive driving licence, is a barrier to this type of ridesharing. In some countries it is prohibited. As an example, in Italy, UberPool and UberX are prohibited, while UberPremium is permitted because it is recognised the status of 'rent with driver' service ('noleggio con conducente') in compliance with existing legislation. Moreover, several studies show that TNCs may increase congestion since they decrease, *de facto*, car occupancy (see more below).

A reduction of vehicle-kilometres travelled, with benefits in terms of congestion and of environmental and safety externalities, is expected from ridesharing. However, the empirical evidence contradicts this expectation. A study by Schaller (2021) on Uber and Lyft services in four cities in the United States and in suburban areas of California showed that ridesharing led to at least a doubling of vehicle-kilometres travelled when comparing ridesharing trips with users' previous mode. This is mainly due to addition of dead-head kilometres before each pick-up and to travellers switching to ridesharing from public transport, biking and walking. Another study by Diao et. al. (2021) examined the impacts of TNCs in the United States. It found that TNCs increased road congestion while transit ridership declined with an insignificant change in vehicle ownership.

The evidence from simulation studies is more ambiguous because it highly depends on the assumptions made and on the city structure. As an example, Caulfield (2009) estimated a significant amount of CO_2 savings from ridesharing in Dublin, when respondents were to rideshare for a return trip, 5 days per-week, 44 weeks a year. Ridesharing reduced annual CO_2 emissions by 12 674 t with respect to travelling alone by car, when only 4% of the respondents rideshare to work. In a simulation study conducted over 247 cities worldwide, Tikoudis et al. (2021) found that in public transport dependent cities ridesharing may draw users away from this mode contributing to a net increase in CO_2 emissions. In car dependent urban areas, strong preference for private car and low density may attenuate the ridesharing environmental impact mitigation potential. Beojone and Geroliminis (2021) found longer travel times in Shenzhen, China, because of higher empty kilometres travelled by idle vehicles without assigned passengers.

These findings make a case for peer-to-peer ridesharing, since it should suffer to a lower extent from empty-vehicle travel. Peer-to-peer ridesharing in cities is the focus of the current paper. There are two key service features. The first is matching between demand, i.e. passengers, and supply, i.e. drivers.

The second is pricing. This refers to the transaction price that is paid by passengers to drivers. Both are determined with the support of a digital platform. In the transportation literature, a number of authors have adopted an equilibrium approach. At equilibrium, no agent (passenger or driver) has an incentive to change choice. Choices include travelling alone versus ridesharing, matching and route. Matching and pricing are computed at equilibrium. An up-to-date review of the relevant literature is found in Wang et al. (2021).

Most recently, Chen and Di (2021) have proposed a bi-level network design approach, where pricing is optimized under the equilibrium constraints for matching and route choices. Two criteria are considered for optimal prices: maximization of social welfare and maximization of platform profit. The platform profit criterion is typical of the literature from production economics, which, however, does not consider the network dimension of the problem. See, among the others, Özkan (2020) and Dong and Leng (2021). The latter reference provides an up-to-date review of this literature.

No contribution is found where matching is optimised with reference to a social objective in a static model. The paper aims to fill this gap. Indeed, peer-to-peer ridesharing appears particularly promising when a social planner manages the platform and matching between passengers and drivers minimizes the sum of travel costs (travel time and fuel) and environmental costs. The paper evaluates the cost savings brought about by a socially optimal ridesharing scheme on the baseline scenario where all travel alone by car or public transport. Additionally, the interest is in the distribution analysis to assess how the ridesharing scheme impacts the different agents (passengers, drivers and social planner). Pricing decisions are determined with the aim of guaranteeing that both passengers and drivers have the right incentives to participate in the ridesharing scheme.

Socially optimal ridesharing has been analysed in the stylised case of a commuting corridor with the dynamic bottleneck model (de Palma et al., 2020). We extend the analysis to networks of general topology with consideration of traveller's route choice only. Our framework is static because, contrarily from the bottleneck model, departure time choices are not modelled. The paper does not exclude that the network is congested, but travel times are considered to be unaffected by ridesharing. This is clearly an approximation. Nevertheless, the analysis will provide novel and non-intuitive insights.

Our aim is to build on the literature about the economics of matching. This has investigated, in particular, college admission, marriage, and the job market where workers are matched with firms. A classical reference is Roth and Sotomayor (1990). Galichon (2018) provides a comprehensive treatment of the mathematical models of the job market. This market has commonalities with ridesharing because demand of different types (firms and passengers with different origin and destination) is matched with supply of different types (workers and drivers with different origin and destination) and a price is paid by the buyers (firms and passengers) to the sellers (workers and drivers). An environmental externality in the matching problem has not been explored so far in the literature. This leads to a new type of problem: social optimal matching with externalities. In this approach we assume that travel times are fixed, i.e. that the level of congestion is fixed.

The optimal matching problem will be formulated as a linear programming problem. Matching flows are the decision variables, which are required to be integer. However, we will prove that there is no need to formulate an integer linear programming problem, because an integer solution exists to the continuous problem and is provided by the commonly used simplex algorithm.

A key concept in the literature is matching stability: no pair of worker and firm prefers each other to any of the current partners. This is a state of equilibrium which does not require the action of a social planner. The paper will apply stability to ridesharing. A key finding in the absence of externalities is the stability of the matchings that are dictated by a social planner who maximizes the sum of the payoffs of worker and firm from matching. The paper will show that this finding applies to ridesharing, when congestion and other externalities are not considered. Clearly, different matchings are obtained when the social planner minimizes not only the travel costs of passengers and drivers, but environmental costs as well. An analysis of the differences between the case with environmental costs in the objective function of the social planner and the case without will be provided.

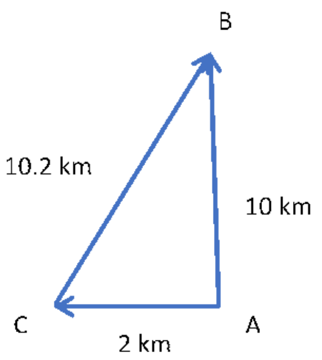
The remainder of the paper is as follows. The ridesharing scheme and associated matching and pricing are introduced in Section 2, and illustrated with a worked example related to a simple triangle network. Section 3 includes the formulation of the socially optimal matching problem along with the equivalence with stable matching. Section 4 provides the numerical analysis of the socially optimal ridesharing scheme in the case of the Sioux Falls network. Finally, the results and the directions of future work are discussed in Section 5.

2. The scheme and a simple example

The socially optimal ridesharing scheme with incentives to participate is as follows. There is a fixed number of subscribers, passengers and drivers, to the ridesharing scheme. Part of passengers would travel alone by car, part would travel alone by public transport (PT). A driver can be alone or with a passenger in the car. The social planner identifies the socially optimal matchings. Individual travel cost (travel time and fuel cost) and emission cost are relevant in this regard. All subscribers conform to the socially optimal matchings dictated by the social planner. Each driver is matched with only one passenger and each passenger with only one driver. Travelling alone can be socially optimal. Matching is termed balanced if we have the same total number of passengers and drivers, unbalanced otherwise. Both cases are considered. Since there is very little empirical literature on estimation of inconvenience cost, we prefer to ignore it rather than making ad-hoc hypotheses. We assume that travel time is constant and focus our attention on the environmental externality.

The social planner evaluates the maximum willingness to pay of each passenger (it is equal to the cost saving when matching, with respect to travelling alone), and the minimum willingness to accept of each driver (it is equal to the cost increase when matching, with respect to travelling alone). Travel time cost, fuel cost and PT fare are relevant in this regard. The social planner receives the maximum willingness to pay from each passenger and pays the minimum willingness to accept to each driver. With this mechanism, neither passenger nor driver loses. The social planner gains the difference between what is paid by the passenger and what is received by the driver. This difference is never negative because there is always the possibility of travelling alone. Additionally, a modified mechanism can be designed where every passenger and every driver gains, if each passenger pays slightly less than the maximum willingness to pay and each driver receives slightly more than the minimum willingness to accept. In this case, there can be matchings with a deficit for the social planner which need cross-subsidisation. Post-processing evaluates whether the social planner overall budget constraint is satisfied, meaning that a net surplus exists, or not. In the latter case, subsidies to the scheme are needed.

A simple example below illustrates the socially optimal ridesharing scheme from a simple triangle network example. Arc lengths (km) are provided for the A-B-C network in Figure 1. Let A, B and C be origin and destination nodes.



OD flows

	A	B	C
A	0	1	0
B	0	0	0
C	0	1	0

Fig.1. Triangle network example

There are two travellers on the network: one passenger goes from C to B and one driver from A to B. Travel speed on the links is 40 km/h, value of travel time is 12.96 €/h, and fuel cost is 0.16 €/km (fuel consumption 10km/l and fuel price is 1.6 €/l). Environmental cost is calculated for CO_2 emissions. Average emission rate is 120 g CO_2/vkm for cars (average value in Europe) and it costs 100 €/t CO_2, so the emission cost is 0.01 €/vkm (It is 1.14 €-cent/vkm in the Handbook on the External Costs of Transport by the European Commission, 2019). Walking speed is 3.6 km/h and value of travel time for walking is 4.32 €/h. Table 1 shows values of the different cost items by arc.

Table 1. Cost components by arc in triangle network

Arc	Travel time (minutes)	Travel time cost (€)	Fuel cost (€)	Emission cost (€)
A-B	15	3.15	1.6	0.114
A-C	3	0.63	0.32	0.03
C-B	15.3	3.213	1.632	0.116

First, a baseline scenario is considered, where everyone travels alone and the costs are computed for both travellers on link A-B and C-B. The social planner will match passenger C-B to the driver A-B by minimizing the matching cost. To find the optimal match, travel cost (travel time and fuel) and emission cost are considered. At optimum, driver will do detour from A to C and will match with passenger C-B. The cost of matching and savings when moving from the baseline scenario to the socially optimal matching scenario are provided in Table 2.

In the baseline scenario both passenger and driver will have travel time, fuel and emission costs. When they match to share the ride, the fuel and emission costs are only applicable to driver for her trip A-C-B including detour A-C. Society has a total saving of 0.671 € in terms of total social cost, equal to the difference between cost in the baseline scenario and cost in the socially optimal matching scenario. This simple network explains the benefit of socially optimal ridesharing that we explore hereafter for larger networks.

Passenger has a maximum willingness to pay of 1.632 € and driver has a minimum willingness to accept of 1.045 €. If passenger pays 1.632 € and driver receives 1.045 €, then neither passenger nor driver loses, and social planner ends up with a profit of 0.587 €.

Table 2. Traveller's cost in triangle network for the baseline and the socially optimal matching scenario

		Baseline scenario					Socially optimal matching scenario				
	OD	Route	Travel time cost (€)	Fuel Cost (€)	Emission cost (€)	Total Cost (€)	Route	Travel time cost (€)	Fuel cost (€)	Emission cost (€)	Total cost (€)
Passenger	C-B	C-B	3.213	1.632	0.116	4.961	C-B	3.213	0	0	3.213
Driver	A-B	A-B	3.15	1.6	0.114	4.864	A-C-B	3.843	1.952	0.146	5.941
Society			6.363	3.232	0.23	9.825		7.056	1.952	0.146	9.154

3. Mathematical formulation of socially optimal ridesharing

In this section, we look into socially optimal ridesharing, in which the social planner determines matchings by minimizing the overall social cost.

3.1 Notation

Indexes and superscripts

a, A	arc and car index
B	PT (Public Transport) index
D	driver superscript
i, j	passenger and driver OD pair index
P	passenger superscript
W	walk index

Parameters and fixed entities

c_{Ai}, c_{Bi}	social cost of one user travelling alone by car and by PT on OD pair i [€/user]
c_{ij}	social cost of matching a passenger of OD pair i with a driver of OD pair j [€/user]
k_{1a}	cost per vehicle-km of fuel consumption on arc a (speed dependent) [€/vkm]
k_{2a}	cost per vehicle-km of emissions on arc a (speed dependent) [€/vkm]
L_a	length of arc a [km]
M	number of OD pairs
N_i^D	number of drivers of OD pair i [user]
N_{Ai}^P, N_{Bi}^P	number of car and PT passengers of OD pair i in the baseline scenario [user]

r_{Aij},r_{Bij}	reservation price of passengers (maximum willingness to pay) who travel by car and by PT in the baseline scenario associated with matching ij [€/user]
r_{Dij}	reservation price of drivers (minimum willingness to accept) associated with matching ij [€/user]
R_{Ai},R_{Bi}	user optimal travel route by car and by PT of OD pair i
R_{ij}^{D}	socially optimal driver's travel route when matching OD pair i with OD pair j (optimum of four cases)
R_{ij}^{P}	socially optimal passenger's travel route when matching OD pair i with OD pair j (optimum of four cases)
R_{Aij}^{P}	socially optimal passenger's route travelled by car when matching OD pair i with OD pair j (optimum of four cases)
R_{Wij}^{P}	socially optimal passenger's route travelled walking when matching OD pair i with OD pair j (optimum of four cases)
t_{Aa}, t_{Ba},t_{Wa}	travel time by car, by PT and by walk on arc a [hour]
α_A,α_B,α_W	value of travel time when using a car, PT and when walking [€/hour]
γ	gain factor
φ	PT fare per trip [€/user]

Decision variables

x_{ij}	number of passengers of OD pair i matched with drivers of OD pair j [user]
x_{Aij}	number of passengers who travel by car in the baseline scenario of OD pair i and who are matched with drivers of OD pair j [user]
x_{Bij}	number of passengers who travel by PT in the baseline scenario of OD pair i and who are matched with drivers of OD pair j [user]

Other variables

p_{ij}	transfer payment in matching ij [€/user]
S_{ij}	social planner cost for matching ij [€]
T_{ij}	social planner revenue for matching ij [€]
u_{Ai}	surplus of passenger (reservation price minus transfer payment) of OD pair i who travels by car in the baseline scenario [€/user]
u_{Bi}	surplus of passenger (reservation price minus transfer payment) of OD pair i who travels by PT in the baseline scenario [€/user]
v_j	surplus of driver (transfer payment minus reservation price) of OD pair j [€/user]

3.2 Social costs

Matching cost

The social cost associated with matching of one passenger of OD pair i with one driver of OD pair j is given by the cost of passenger's travel time, plus the cost of driver's travel time, plus fuel consumption cost, plus environmental cost:

$$c_{ij} = \sum_{a\in R^P_{wij}} \alpha_w t_{wa} + \sum_{a\in R^P_{Aij}} \alpha_A t_{Aa} + \sum_{a\in R^D_{ij}} \alpha_A t_{Aa} + \sum_{a\in R^D_{ij}} k_{1a} L_a + \sum_{a\in R^D_{ij}} k_{2a} L_a. \qquad (1)$$

Notice that fuel cost and environmental cost are assumed proportional to vehicle-km travelled. Environmental costs account for emissions of CO_2 and local pollutants.

Cost of travelling alone

The social cost of one user travelling by car is given by the cost of travel time plus fuel cost plus environmental cost:

$$c_{Ai} = \sum_{a\in R_{Ai}} \alpha_A t_{Aa} + \sum_{a\in R_{Ai}} k_{1a} L_a + \sum_{a\in R_{Ai}} k_{2a} L_a. \qquad (2)$$

The social cost of one user travelling by PT is given by the cost of travel time:

$$c_{Bi} = \sum_{a\in R_{Bi}} \alpha_B t_{Ba}. \qquad (3)$$

A situation of under capacity is assumed for PT. This means that the marginal user does not add to vehicle-km and, therefore, fuel costs and environmental costs are omitted.

3.3 Optimal routes

A user's route is the route the individual selects to her destination. The optimal route in matching is the route that twins a passenger with a driver minimising social cost for given origin and destination.

Optimal routes associated with matching

Optimal route for matching is identified as soon as the social planner decides who will do detour between driver and passenger at both origin and destination. Matching in our case will happen either at drivers' origin or at passengers' origin and the same applies for the destination. Four matching cases are considered to take all possibilities into account (Fig. 2). In case 1: passenger walks to driver origin then both travels together to driver destination and then again passenger walks to her destination. In case 2: passenger detours at origin and driver detours at destination. In case 3: driver detours at origin and passenger detours at destination. In case 4: driver detours at both origin and destination.

Routes $R_{ij}^{P} = R_{Wij}^{P} \cup R_{Aij}^{P}$ chosen by passengers when matched, and routes R_{ij}^{D} chosen by drivers when matched, are identified off-line based on the optimum, i.e., minimum social cost c_{ij}, of the four cases illustrated in the figure. It is implicitly assumed that vehicles are guided in their route choice.

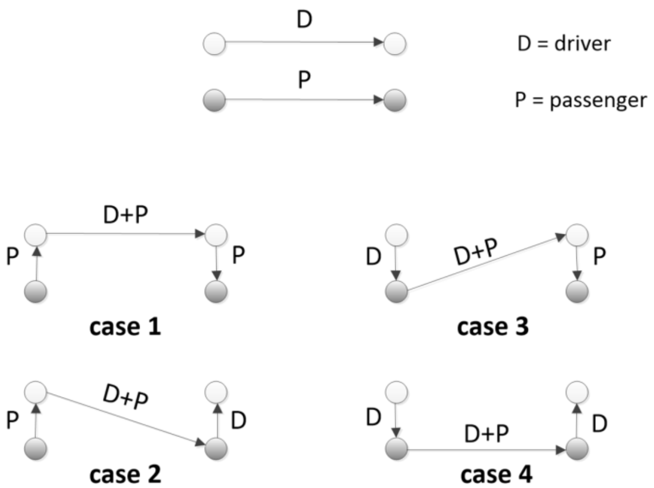


Fig. 2. The four matching cases of driver to passenger, where either passenger or driver detours at origin and destination

Here, just one passenger matching is evaluated to keep the formulation simple. It is straightforward to extend to multiple matchings i.e., matchings involving several driver-passenger pairs.

Optimal routes associated with travelling alone

The routes R_{Ai} travelled by users by car and the routes R_{Bi} travelled by users by PT when matching is not implemented are identified based on minimum travel time cost. The shortest route cost considers only travel time cost and not social cost.

3.4 Socially optimal matching

We formulate the socially optimal matching problem in the general case, where travelling alone can be socially optimal, and where the total number of passengers can be higher than, equal to or lower than the total number of drivers. There are M OD pairs that are identical for passengers and drivers.

We introduce a dummy passenger OD pair denoted by the index $M + 1$. If a driver is assigned to this dummy passenger OD pair, then it means that the driver is not matched and travels alone.

We introduce a dummy driver OD pair denoted by the index $M + 1$. If a passenger is assigned to this dummy driver OD pair, then it means that the passenger is not matched and travels alone by car or by PT as in the baseline scenario.

The decision variables are:

$x_{ij}, i, j = 1, \dots, M$ flow of passengers of OD pair i who are matched with drivers of OD pair j

$x_{Ai,M+1}, i = 1, \dots, M$ flow of passengers of OD pair i who travel by car in the baseline and who are not matched

$x_{Bi,M+1}, i = 1, \dots, M$ flow of passengers of OD pair i who travel by PT in the baseline and who are not matched

$x_{M+1,j}, j = 1, \dots, M$ flow of drivers of OD pair j who are not matched.

Formulation of Problem $P1$ - Socially optimal matching: *the total social cost of matchings and travel alone is minimised subject to the assignment constraints and the non-negativity constraints*

$$\min \sum_{i=1,\dots,M} \sum_{j=1,\dots,M} c_{ij} x_{ij} + \sum_{i=1,\dots M} c_{Ai} x_{Ai,M+1} + \sum_{i=1,\dots M} c_{Bi} x_{Bi,M+1} + \sum_{j=1,\dots,M} c_{Aj} x_{M+1,j}, \tag{4}$$

subject to:

$$\sum_{j=1,\dots,M} x_{ij} + x_{Ai,M+1} + x_{Bi,M+1} = N_{Ai}^P + N_{Bi}^P,\; x_{Ai,M+1} \le N_{Ai}^P, x_{Bi,M+1} \le N_{Bi}^P,\; i = 1, \dots, M, \tag{5}$$

$$\sum_{i=1,\dots,M} x_{ij} + x_{M+1,j} = N_j^D, \quad j = 1, \dots, M, \tag{6}$$

$$x_{ij} \ge 0,\; i = 1, \dots, M; j = 1, \dots, M, \tag{7a}$$

$$x_{Ai,M+1} \ge 0,\; x_{Bi,M+1} \ge 0, \quad i = 1, \dots M, \tag{7b}$$

$$x_{M+1,j} \ge 0, \quad j = 1, \dots, M. \tag{7c}$$

The objective function in Eq. (4) equals the social cost of all matchings plus the social cost of not matching passengers who travel by car in the baseline, plus the social cost of not matching passengers who travel by PT in the baseline, plus the social cost of not matching drivers.

The first set of constraints of Eqs (5) is the assignment of passengers of each OD pair to drivers (true drivers and dummy driver). The second set of constraints of Eqs (6) is the assignment of drivers of each OD pair to passengers (true passengers and dummy passenger). The third set of constraints of Eqs (7a), (7b) and (7c) is non-negativity of decision variables.

An existence theorem is provided in Proposition 1 below.

The optimal solution allocates each total passenger OD flow to the different true driver od pairs. The solution is indeterminate in terms of optimal allocation of passenger flows who used car in the baseline scenario, and of passenger flows who used PT in the baseline scenario. Therefore, we introduce the additional decision variables:

$x_{Aij}, i, j = 1, \dots, M$ flow of passengers of OD pair i who travel by car in the baseline and who are matched with driver OD pair j

$x_{Bij}, i, j = 1, \dots, M$ flow of passengers of OD pair i who travel by PT in the baseline and who are matched with driver OD pair j.

The reason why we have not formulated the problem in terms of these decision variables is indeterminacy: the matching cost in the objective function depends on the total matching flow $x_{ij} = x_{Aij} + x_{Bij},\ i, j = 1, \dots, M$. To find the values of $x_{Aij}, x_{Bij}, i, j = 1, \dots, M$, we maximise the profit of the social planner.

The social planner matches drivers and passengers, and for each matching passengers have some willingness to pay and drivers have some willingness to accept. The maximum willingness to pay for the passenger is the difference between the cost travelling alone and the cost ridesharing. The extra money she will save by sharing a ride will be equal to her maximum willingness to pay. Since the passenger is the buyer, her maximum willingness to pay is her reservation price. The minimum willingness to accept for the driver is the difference between the cost ridesharing and the cost travelling alone. The extra money she will pay when sharing a ride will be equal to her minimum willingness to accept. Since the driver is the seller, her minimum willingness to accept is her reservation price. If passenger or driver travels alone, then her reservation price is zero.

Given the matching $ij, i, j = 1, \dots, M$, the following hold.

The reservation price of a passenger (maximum willingness to pay) of od pair i travelling alone by car is the difference between the cost of car passenger travelling alone (travel time cost by car and fuel cost) and cost of ridesharing (travel time cost walking and by car):

$$r_{Aij} = \sum_{a \in R_{Ai}} \alpha_A t_{Aa} + \sum_{a \in R_{Ai}} k_{1a} L_a - \left(\sum_{a \in R_{wij}^P} \alpha_w t_{wa} + \sum_{a \in R_{Aij}^P} \alpha_A t_{Aa} \right). \tag{8}$$

The reservation price of a passenger (maximum willingness to pay) of od pair i travelling alone by PT is the difference between the cost of PT passenger travelling alone (travel time cost by PT and trip fare) and cost of ridesharing (travel time cost walking and by car):

$$r_{Bij} = \sum_{a \in R_{Bi}} \alpha_B t_{Ba} + \varphi - \left(\sum_{a \in R_{wij}^P} \alpha_w t_{wa} + \sum_{a \in R_{Aij}^P} \alpha_A t_{Aa} \right). \tag{9}$$

The reservation price of a driver (minimum willingness to accept) of od pair j is: the difference between the cost of driver ridesharing (travel time cost by car and fuel cost when matched) and cost of travelling alone (travel time cost by car and fuel cost).

$$r_{Dij} = \sum_{a\in R_{ij}^{D}} \alpha_A t_{Aa} + \sum_{a\in R_{ij}^{D}} k_{1a} L_a - \Big(\sum_{a\in R_{Aj}} \alpha_A t_{Aa} + \sum_{a\in R_{Aj}} k_{1a} L_a \Big). \quad (10)$$

Assume that the social planner receives from the passenger her maximum willingness to pay and pays to the driver her minimum willingness to accept. In that case, the passenger and the driver are neutral, i.e. they incur neither gain nor loss. The profit of the social planner equals the difference between the maximum willingness to pay of the passenger and the minimum willingness to accept of the driver. If the profit is positive, then the social planner will have a surplus. If the profit is negative, then the social planner will have a deficit. However, it is to see that, since there is the option of travelling alone, the profit is never negative. The sum of profits of all matchings equals the total social planner profit.

More generally, assume that the passenger and the driver are neutral or gain. Let $\gamma \geq 0$ be the gain factor. It measures the fraction of maximum willingness to pay that is not paid by the passenger. It also measures the fraction of minimum willingness to accept that is received in addition by the driver. If $\gamma = 0$ there is no gain for passengers and drivers (the previous neutrality case). The higher the γ value we consider, the higher the gain for drivers and passengers and the lower the budget available for the social planner. This is because the passenger pays less than her maximum willingness to pay and the driver receives more than her minimum willingness to accept.

For matching ij, the social planner revenue is the sum of the payments by passengers who would travel alone by car plus the payments by passengers who would travel alone by PT:

$$T_{ij} = (1-\gamma) r_{Aij} x_{Aij} + (1-\gamma) r_{Bij} x_{Bij}, i,j = 1, \dots, M, \quad (11)$$

while the social planner cost is the sum of the payments to drivers:

$$S_{ij} = (1+\gamma) r_{Dij} x_{ij}, i,j = 1, \dots, M. \quad (12)$$

Formulation of Problem $P2$ - Socially optimal matching step 2: *the social planner's profit, equal to the difference between revenues and costs, is maximised subject to the demand conservation constraints and the non-negativity constraints*

$$\max \sum_{i=1}^{M} \sum_{j=1}^{M} \big(T_{ij} - S_{ij}\big), \quad (13)$$

subject to:

$$x_{Aij} + x_{Bij} = x_{ij}, i,j = 1, \dots, M, \quad (14)$$

$$\sum_{j=1,\dots,M} x_{Aij} + x_{Ai,M+1} = N_{Ai}^{P}, \quad i = 1, \dots, M, \quad (15a)$$

$$\sum_{j=1,\dots,M} x_{Bij} + x_{Bi,M+1} = N_{Bi}^{P}, \quad i = 1, \dots, M, \quad (15b)$$

$$x_{Aij} \geq 0, x_{Bij} \geq 0,\ i = 1, \dots, M; j = 1, \dots, M. \tag{16}$$

The objective function in Eq. (13) is the total profit of the social planner. The first set of constraints of Eqs (14) is the flow conservation for each matching. The second set of constraints of Eqs (15a) and (15b) is the flow conservation for each passenger OD pair. The third set of constraints of Eqs (16) is non-negativity of decision variables.

The two optimization Problems $P1$ and $P2$ need to be solved in two steps in series. Both optimization problems are linear. The following proposition justifies why integrality constraints on the flow decision variables are un-necessary.

Proposition 1. *Consider Problems $P1$ and $P2$. If the constants of the constraints are integer, then there exists one optimal integer solution to each problem. Solution to Problems $P1$ and $P2$ provided by the simplex algorithm are integer.*

Proof. First, we need to re-formulate the constraints of both problems in compact form as $Ax \leq b, x \geq 0$. Then, in the light of Farkas lemma, a solvability theorem for a finite system of linear inequalities (Gale et al., 1951), a feasible solution exists. The constraints define a polyhedron which is not unbounded in the direction of the gradient of the objective functions. Therefore, an optimal solution exists. Additionally, the matrix A is totally unimodular, in the light of theorem 5.24 in Korte and Vygen (2008), because there is a partition of the lines of matrix A such that, for each column, the sum of the elements in the first partition minus the sum of the elements in the second partition is equal to either 0, 1 or -1.

As a consequence of total unimodularity of A and integrality of b, in the light of the Hoffman and Kruskal theorem (corollary 19.2a in Schrijver, 1998), the polyhedron defined by the constraints has vertices with all integer coordinates. This proves that an optimal integer solution exists, because there exists an optimal solution which is at a vertex, by the maximum principle for convex functions (theorem 3.10.11 in Niculescu and Persson, 2018). Finally, the simplex algorithm provides optimal integer solutions, because the algorithm provides optimal solutions at vertices of the polyhedron (Dasgupta et al., 2008). □

Uniqueness of the solution, though, is not guaranteed. This implies that we can obtain the same optimal objective function for different values of the decision variables.

3.5 Relationship between socially optimal matching and stability

This section is based on the assignment model of workers to firms found in chapter 3 in Galichon (2018) and chapter 8 in Roth and Sotomayor (1990). Galichon considers the case where workers of one type are matched with firms of more than one type. He does not allow unmatched agents. In contrast, Roth and Sotomayor (1990) considers the case of one-to-one matching. They allow for unmatched agents. Our setting needs adaptation of both sources, since we deal with the case where passengers of the same type are matched with drivers of different types and unmatched agents are allowed.

Consider a setting where passengers are matched with drivers and a price is paid by each passenger to the matched driver. Assume that passenger of OD pair i who travels by car or by PT in the baseline is matched with driver of OD pair j. The surplus u_{Ai} or u_{Bi} of passenger from the matching is the difference $r_{Aij} - p_{ij}$ or $r_{Bij} - p_{ij}$ between her reservation price and the price actually paid. The surplus v_j of the driver is the difference $p_{ij} - r_{Dij}$ between the price actually paid and her reservation price. The profit of the matching is the sum of passenger surplus plus driver surplus. This equals the difference $r_{Aij} - r_{Dij}$ or $r_{Bij} - r_{Dij}$ between the reservation prices of the two agents.

Notice that if passenger travels alone, then $u_{Ai} = 0$ or $u_{Bi} = 0$. Similarly, if driver travels alone, then $v_j = 0$. Notice also that if passengers who travel by car in the baseline of a given OD pair are matched with drivers of different OD pairs, then they gain the same surplus while the price actually paid is differentiated by OD pair. The same holds for passengers who travel by PT in the baseline. Symmetrically, if drivers of a given OD pair are matched with passengers of different OD pairs, then they gain the same surplus while the price actually paid is differentiated by OD pair.

Definition 1 (outcome). *An outcome is the specification* (x, u, v) *of the matching flow for each pair* ij *including travelling alone, of the passenger surplus for each true passenger* i*, and of the driver surplus for each true driver* j*.*

Definition 2 (outcome feasibility). *An outcome* (x, u, v) *is feasible if the total profit generated from matching is equal to the total quantity of profit redistributed to passengers and drivers:*

$$\sum_{i=1}^{M} \sum_{j=1}^{M} \left[(r_{Aij} - r_{Dij}) x_{Aij} + (r_{Bij} - r_{Dij}) x_{Bij} \right] = \sum_{i=1}^{M} \left(u_{Ai} N_{Ai}^{P} + u_{Bi} N_{Bi}^{P} \right) + \sum_{j=1}^{M} v_j N_i^{D}. \tag{17}$$

Eq. (17) shows how the total profit generated at the pairwise level (left-hand side) is redistributed at the individual level (right-hand side).

Definition 3 (outcome stability). *A feasible outcome* (x, u, v) *is stable if the following conditions are satisfied:*

$$u_{Ai}, u_{Bi} \geq 0, i = 1, \ldots, M; \; v_j \geq 0, j = 1, \ldots, M, \tag{18}$$

$$u_{Ai} + v_j \geq r_{Aij} - r_{Dij}, u_{Bi} + v_j \geq r_{Bij} - r_{Dij}, i, j = 1, \ldots, M. \tag{19}$$

Conditions in Eqs (19) express the absence of a blocking pair. Explanation is as follows. Consider an outcome (x, u, v) and assume that there is a passenger of OD pair i and a driver of OD pair j such that $u_{Ai} + v_j < r_{Aij} - r_{Dij}$ or $u_{Bi} + v_j < r_{Bij} - r_{Dij}$. Then, this passenger and this driver will have an incentive to break the current matching (say with, respectively, driver l, and with passenger m) and match together. Such a situation would form a blocking pair which we want to rule out. Notice that inequalities (19) should hold for every pair ij not only for the matched pair.

On the basis of the definition of stability, an outcome is stable if and only if no driver and no passenger has an incentive to change the current matching and match together. It is a concept of equilibrium. We now state the main proposition which sets the relationship between social optimum without environmental costs and outcome stability.

Proposition 2. *An outcome (x, u, v) is stable if and only if the matching flow pattern x is socially optimal for Problem $P1$ without environmental costs. Additionally, if at social optimum without environmental costs $x_{Aij} > 0$ or $x_{Bij} > 0, i, j = 1, \dots, M$, then the surplus of passenger and driver is equal to the profit from matching:*

$$u_{Ai} + v_j = r_{Aij} - r_{Dij}, \tag{20a}$$

$$u_{Bi} + v_j = r_{Bij} - r_{Dij}. \tag{20b}$$

Proof. First, we show that our socially optimal matching Problem $P1$ without environmental costs can be formulated as the problem where the total profit from matchings is maximized. Then, we introduce the dual problem and use some duality theorems.

Consider Problem $P1$ without environmental costs. Using Eqs (11), we introduce the decision variables x_{Aij} and $x_{Bij}, i, j = 1, \dots, M$ and replace constraints (5) with constraints (15a) and (15b). This problem is equivalent to the maximization problem with objective function equal to the difference between the costs perceived by passengers and drivers in the baseline and the costs perceived by passengers and drivers when matching. Indeed, the costs perceived by passengers and drivers in the baseline are a constant. By simple algebra, we obtain the following socially optimal matching problem without environmental costs termed **Problem $P3$**:

$$\max \sum_{i=1}^{M} \sum_{j=1}^{M} \left(r_{Aij} - r_{Dij}\right) x_{Aij} + \sum_{i=1}^{M} \sum_{j=1}^{M} \left(r_{Bij} - r_{Dij}\right) x_{Bij}, \tag{21}$$

subject to:

$$\sum_{j=1}^{M+1} x_{Aij} = N_{Ai}^{P},\ i = 1, \dots, M, \tag{15a}$$

$$\sum_{j=1}^{M+1} x_{Bij} = N_{Bi}^{P},\ i = 1, \dots, M \tag{15b}$$

$$\sum_{i=1}^{M+1}\left(x_{Aij}+x_{Bij}\right)=N_j^D,\ j=1,\dots,M, \tag{22}$$

$$x_{Aij}\geq 0, x_{Bij}\geq 0\ \ i=1,\dots,M; j=1,\dots,M, \tag{16}$$

$$x_{Ai,M+1}\geq 0,\ \ x_{Bi,M+1}\geq 0\ ,\quad i=1,\dots M, \tag{7b}$$

$$x_{A,M+1,j}\geq 0, x_{B,M+1,j}\geq 0\ \ \ j=1,\dots,M. \tag{23}$$

Problem $P3$ is equivalent to the following problem with inequality constraints, termed **Problem $\boldsymbol{P4}$**:

$$\max\sum_{i=1}^{M}\sum_{j=1}^{M}\left(r_{Aij}-r_{Dij}\right)x_{Aij}+\sum_{i=1}^{M}\sum_{j=1}^{M}\left(r_{Bij}-r_{Dij}\right)x_{Bij}, \tag{24}$$

subject to:

$$\sum_{j=1}^{M}x_{Aij}\leq N_{Ai}^P,\ i=1,\dots,M, \tag{25a}$$

$$\sum_{j=1}^{M}x_{Bij}\leq N_{Bi}^P,\ i=1,\dots,M, \tag{25b}$$

$$\sum_{i=1}^{M}\left(x_{Aij}+x_{Bij}\right)\leq N_j^D,\ j=1,\dots,M, \tag{26}$$

$$x_{Aij}\geq 0, x_{Bij}\geq 0\ \ i=1,\dots,M; j=1,\dots,M. \tag{16}$$

Consider now the dual of Problem $P4$, termed **Problem $\boldsymbol{P5}$**:

$$\min\sum_{i=1}^{M}\left(u_{Ai}N_{Ai}^P+u_{Bi}N_{Bi}^P\right)+\sum_{j=1}^{M}v_jN_i^D, \tag{27}$$

subject to:

$$u_{Ai},u_{Bi}\geq 0, i=1,\dots,M;\ v_j\geq 0, j=1,\dots,M, \tag{18}$$

$$u_{Ai}+v_j\geq r_{Aij}-r_{Dij}, u_{Bi}+v_j\geq r_{Bij}-r_{Dij}, i,j=1,\dots,M. \tag{19}$$

At optimum, the objective function of the primal Problem $P4$ equals the objective function of the dual Problem $P5$ (theorem 4.8 in Eiselt and Sandblom, 2010). This proves the first part of the theorem.

Eqs (20a) and (20b) are a consequence of the complementary slackness conditions which hold at optimum (theorem 4.9 in Eiselt and Sandblom, 2010). □

4. Numerical analysis: Sioux Falls network

4.1 The network

The network of Sioux Falls city is used to implement the mathematical model and investigate the impacts of ridesharing. It was first considered by LeBlanc (1975). The Sioux Falls network is frequently used by researchers for test and implementation of models and algorithms. In this network, there are 24 nodes, 76 links, and 24 zones as shown in Figure 3.

The network includes 24x23 = 552 OD pairs. Drivers and passengers are considered in each zone and they are willing to share a ride. There is a total of 360 600 travellers on the network, 10% of them are considered to be passengers and 10% are considered to be drivers willing to share the ride. Driver and passenger OD flows are generated randomly keeping the sum of flows for each OD equal to 20% of the total flow. The social planner will match 36 060 drivers to 36 060 passengers optimally for ridesharing (balanced case). Travellers are considered to use car and PT in the baseline scenario and the ratio is fixed exogenously. There is a total of 28 838 (80%) passengers travelling by car and 7 222 (20%) by PT.

Free-flow travel times are considered for cars. PT travel times are 50% higher than by car because of lower speed and delay at stops. The following parameter values are assumed: value of time by car is 12.96 €/h, value of time by PT is 18 €/h, value of time walking is 4.32 €/h, fuel cost is 0.16 €/km, cost of emissions, restricted to CO_2, is 0.01 €/vkm. PT fare is 1.5 €/user.

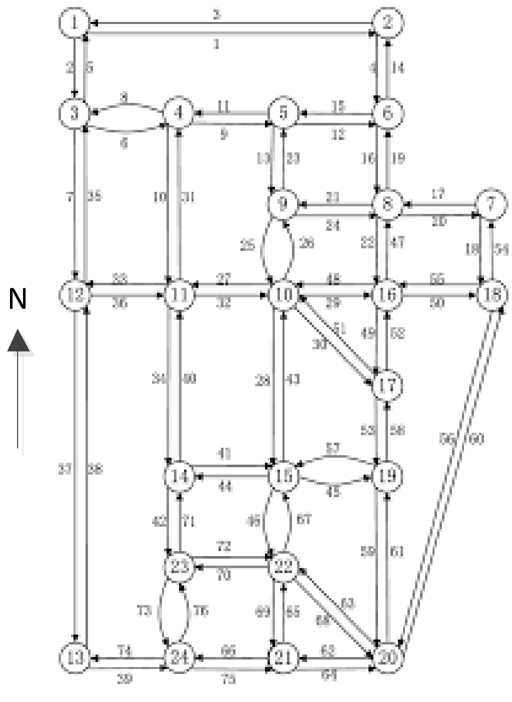


Fig.3. Sioux Falls network

4.2 Results

Optimal matchings result from solutions to Problems $P1$ and $P2$. Solution to Problem $P1$ provides optimal social costs, while social planner's budget analysis is carried out on the basis of solution to Problem $P2$. The code is written in Python language. PuLP library and the simplex algorithm are used for optimization.

Matching with environmental costs

The results are computed here for both the baseline scenario (no ridesharing) and the socially optimal matching scenario with environmental costs in the objective function. First, the shortest route between all the zones is calculated using Dijkstra algorithm. Baseline scenario costs are computed using the travel time cost, fuel cost and environmental cost based on the shortest travel routes. In the baseline scenario there are 72 120 drivers and passengers travelling alone between different OD using car or PT.

In the socially optimal matching scenario, drivers and passengers are allowed to match. Socially optimal costs are computed using the best of the four matching cases described in Figure 2. Then, the optimal matchings between passengers and drivers are computed.

Total cost for the baseline scenario for all users travelling alone (Eq. 2 and 3) is 214 465 €. PT passengers have high travel time cost because their travel times on the network are considered 50% higher than by car and they have comparatively high travel time cost per kilometre. When matched, total travel time cost decrease because now PT passengers are sharing ride with car drivers and have lower travel time cost. Also, the fuel cost and emission cost decrease in the matching scenario due to less vehicle-kilometres travelled on the network. The total cost when drivers and passengers match optimally (Eq. 4) is 175 716 €. The overall effect of ridesharing on cost saving is positive and there is a net saving of 38 749 € (around 18%) on the baseline scenario. Fuel cost in the baseline scenario is 61 106 €, it is reduced to 39 556 € in the matching scenario and it has the greater effect on the overall cost saving.

Total vehicle-kilometres travelled on the network in the baseline scenario are 381 912 vkm, which are reduced to 247 225 vkm in the matching scenario. This helps reduce congestion since there is overall 35% less vehicle-kilometres on the network in the matching scenario. Less vehicle-kilometres travelled also help reduce the CO_2 emissions. At the rate of 120 g CO_2/vkm, the social optimal ridesharing can save a total of 16 162 kg of CO_2. Table 3 and Table 4 explain in detail the different costs for the baseline and the matching scenarios. In the socially optimal matching, we can see that there are no solo PT passengers since they have higher cost of travelling alone. Therefore, when cost is optimized, they match before the car passengers for each OD.

In the baseline scenario, PT passengers have only travel time cost and solo car drivers and passengers have travel time, fuel and emission cost. In the matching scenario, 13% (9480/72120) of the total users do not match, and detour accounts for 8% of the total CO_2 emissions.

We also notice that there is a total of 251 km walking by 172 passengers who match with drivers. Walking time cost is 301 €. The maximum distance a passenger walks to match is 2 km.

Table 3. Travel costs in the baseline scenario (no matching)

	Number of travellers	Travel time cost (€)	Fuel cost (€)	Emission cost (€)	Total cost (€)
Solo drivers	36 060	67 671	34 372	2 148	104 191
Solo car passengers	28 838	53 632	26 734	1 671	82 037
Solo PT passengers	7 222	28 237	0	0	28 237
Society	72 120	149 540	61 106	3 819	214 465

Table 4. Travel costs in the socially optimal matching with environmental costs

	Number of travellers	Travel time cost (€)	Fuel cost (€)	Emission cost (€)	Total cost (€)
Matching	62 640	117 939	31 557	1 972	151 468
Solo drivers	4 740	5 887	2 990	187	9 064
Solo car passengers	4 740	9 862	5 009	313	15 184
Solo PT passengers	0	0	0	0	0
Society	72 120	133 688	39 556	2 472	175 716

From the social planner's perspective, the budget from matching is important since it can be used to operate the system or can be invested elsewhere, as an example to improve public transport. The total profit of social planner is the difference between total revenue and total cost (Eq. 13). Figure 4 shows some results of the social planner's budget analysis.

When the gain factor is zero, the social planner has surplus in all the matchings. Total revenue gain from car matchings is 21 544 €, from PT matching is 25 845 €, and total cost for matching drivers is 524 €. The social planner profit for matching drivers to passengers is 46 864 € and it shows an average profit of 1.5 € per each of the 31 320 matchings. This shows that there is a monetary gain for society from the ridesharing scheme. When we increase the gain factor for drivers and passengers (Eq. 11 and 12), the social planner incurs deficit for some matchings and her profit decreases.

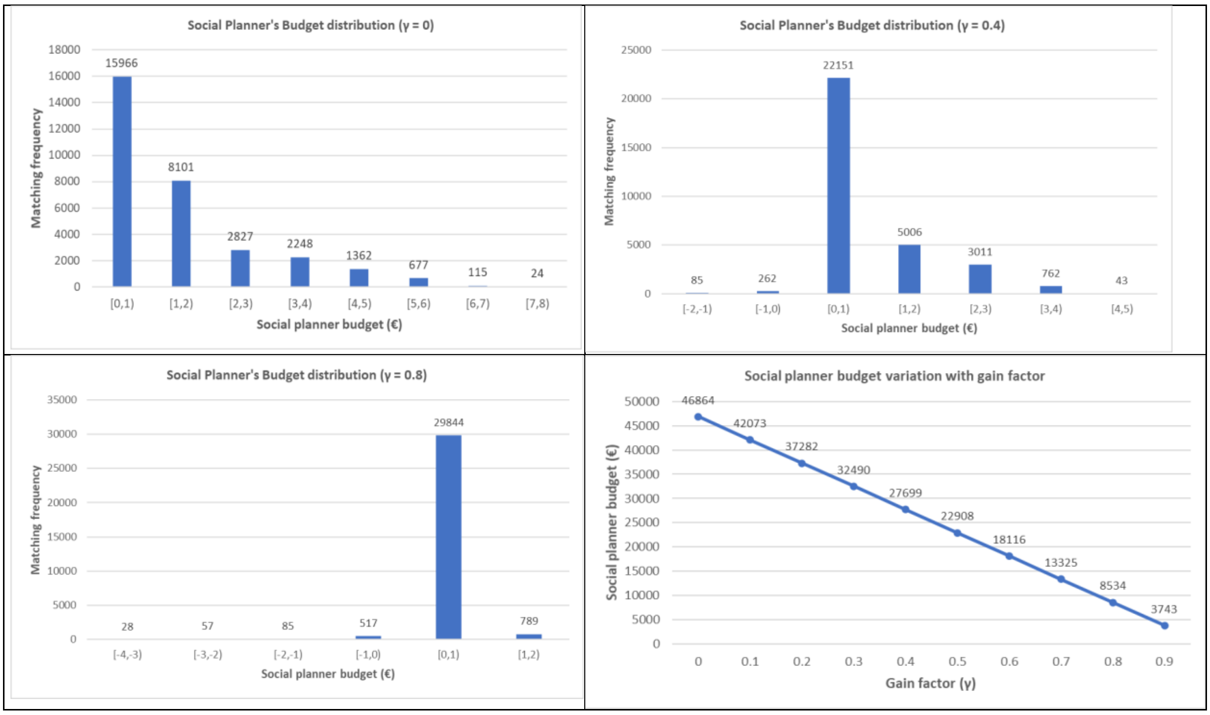


Fig. 4. Socially optimal matching with environmental costs: social planner's budget analysis

Matching without environmental costs

Earlier in the socially optimal matching, we included environmental costs in the objective function. Now, we will check the matching without environmental costs in the objective function and evaluate the difference. The baseline scenario remains the same. Results in the matching scenario in Table 5 show that we have higher total social cost, because we are optimizing the costs without considering environmental costs. The matchings are also less, while fuel and environmental costs increase. Therefore, when passengers are matched with passengers without considering environmental costs, fuel and emission cost savings are comparatively less.

We also notice that there is total of 112 km walking by 96 passengers who match with drivers. Total walking time cost is 134 €. The maximum distance a passenger walks to match is 1.3 km. All values are lower than those in the with environmental costs case.

The results of the social planner's budget analysis are shown in Figure 5. The budget without environmental costs is higher than the budget with environmental costs.

Table 5. Travel costs in the socially optimal matching without environmental costs

	Number of travellers	Travel time cost (€)	Fuel cost (€)	Emission cost (€)	Total Cost (€)
Matching	62 496	1 19 250	32 211	2 014	153 475
Solo drivers	4 812	4 570	2 321	145	7 036
Solo car passengers	4 812	10 106	5 133	321	15 560
Solo PT passengers	0	0	0	0	0
Society	72 120	133 926	39 665	2 480	176 071

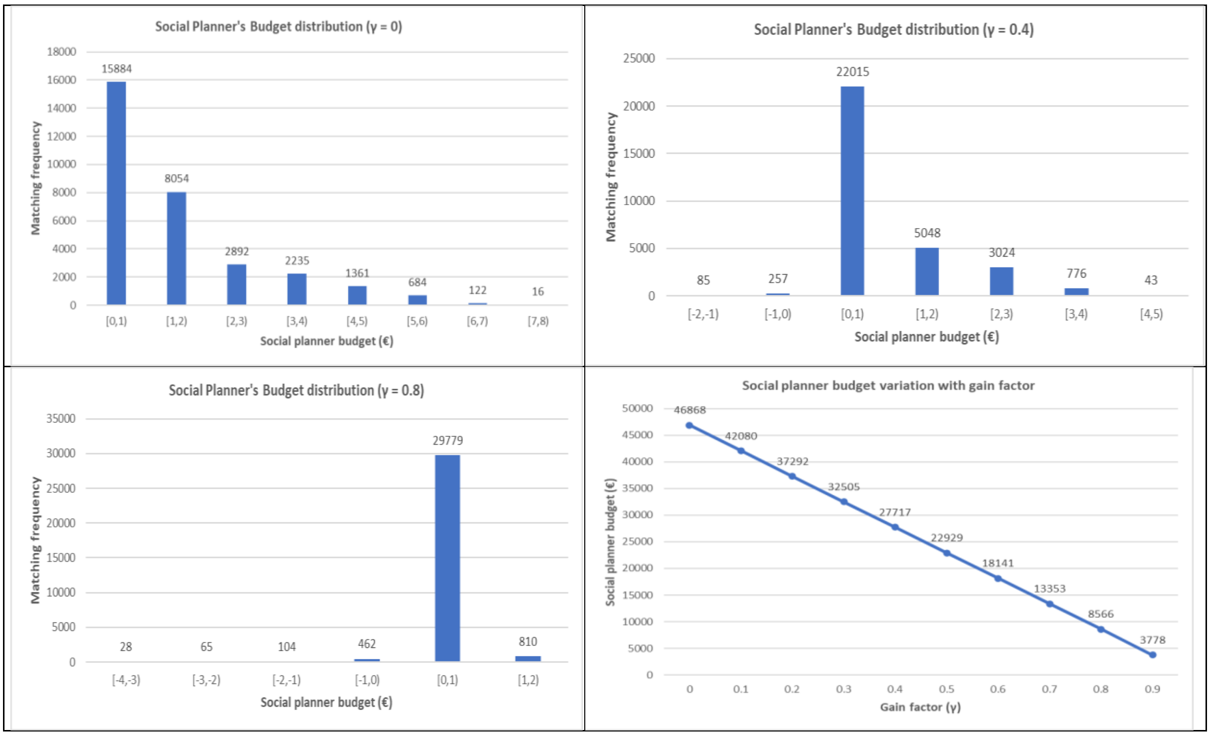


Fig. 5. Socially optimal matching without environmental costs: social planner's budget analysis

Sensitivity analysis with respect to percentage of PT passengers in the baseline scenario

Our analysis until now has been focused on a fixed 20 percentage of PT passengers in the baseline scenario, but when the ratio of PT and car passengers changes, so do the matchings, as well as the social optimal cost and social planner's budget. Sensitivity analysis is addressed to see these variations with respect to change in the percentage of PT passengers. Results are in Figure 6. Cost of travelling alone for PT passengers is high compared to car passengers in the baseline scenario, so, when matched, PT passengers are preferred to match first. When the number of PT passengers is low, they are all matched, and the cost remains constant; when the percentage of PT passengers increases, they also travel alone optimally, and the cost begins to rise.

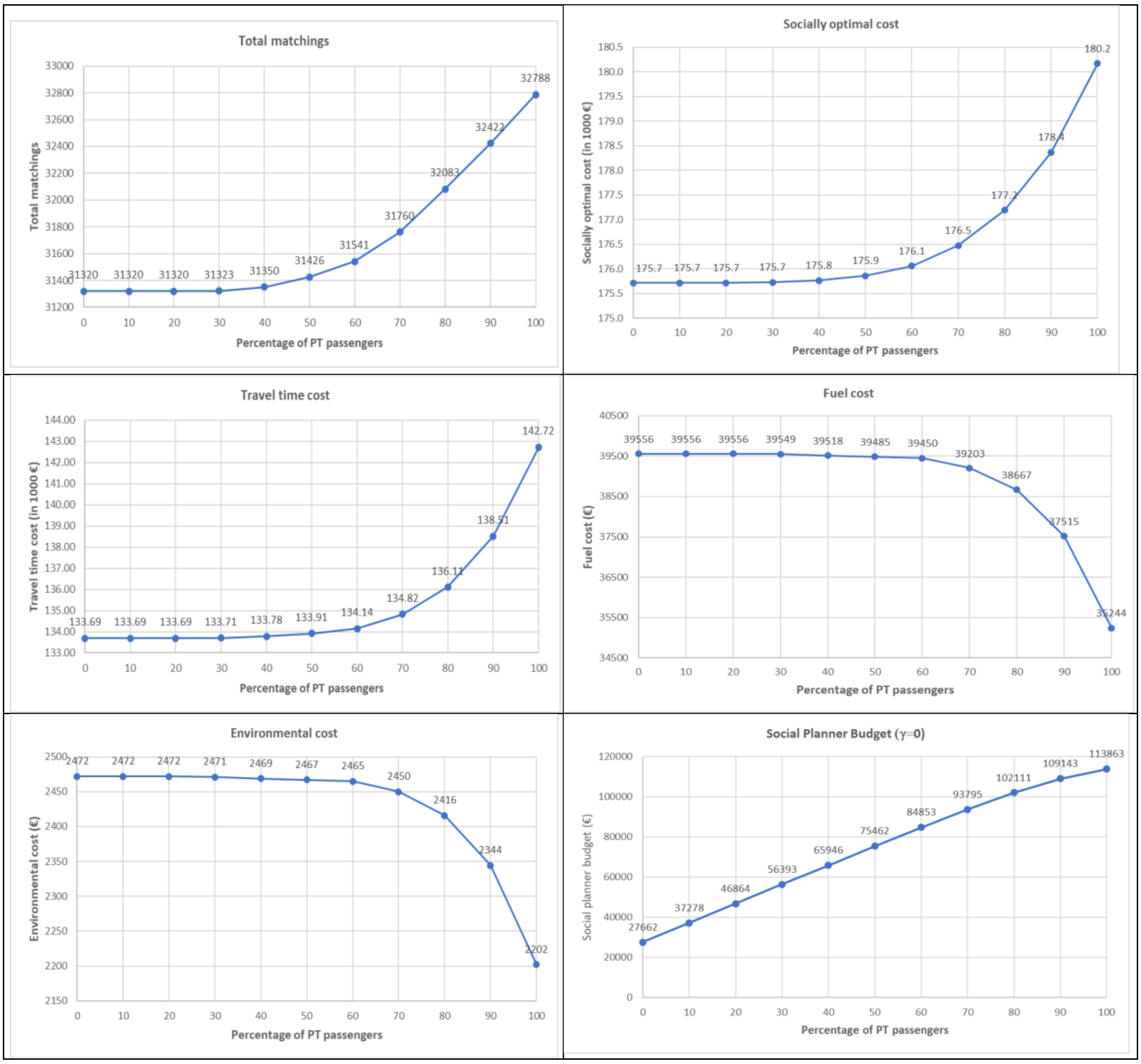


Fig. 6. Sensitivity analysis with respect to the percentage of PT passengers in the baseline scenario

We also obtain more matchings when the share of PT passengers is high, because their cost of travelling alone cost is high. The social planner's budget per matching increases with the percentage of PT passengers.

Matching with externality cost

Externality cost is explained as the externality generated by adding a car on network. Externality cost considered here is equal to the cost of travelling alone. When externality cost is so high, car passengers generate higher externality, thus they are preferred to match because travelling alone will double the externality.

Considering high externality cost helps us understand the situation where other environmental externalities or congestion externality will impact ridesharing. The results in Table 6 reveal a higher number of matchings and lower fuel costs and more car passengers' matchings than PT passengers, implying less vehicle-kilometres travelled on the network.

Table 6. Comparison between the cost of socially optimal matching with environmental costs, without environmental costs and with externality cost

Matching	Number of matchings	Travel time cost (€)	Fuel cost (€)	Emission/exter-nality cost (€)	Total Cost (€)
Socially optimal	31 320	133 688	39 556	2 472	175 716
Without environmental costs	31 248	133 926	39 665	2 480	176 071
With externality cost	33 192	144 936	34 696	103 170	282 802

Impact of variation of passenger OD flow matrix on driver OD flow matrix

The difference between the OD flows of drivers and the OD flows of passengers is critical for matching ease and overall cost saving. Consider the driver OD flow matrix. It is realistic to assume that the passenger OD flow matrix will show a similar OD flow pattern. The interest is in evaluating the impact on the matching pattern, and the associated total social cost, of the variation of the passenger OD flow matrix on the driver OD flow matrix. This analysis can be carried out using the Dirichlet distribution (see Appendix).

The Dirichlet distribution can be used to cut a string into n pieces with different lengths, where each piece has, on average, a designated average length, while allowing some variation in the relative sizes of the pieces. In our case, the average length of one piece is the driver OD flow. From the Dirichlet distribution, 200 passenger OD flow matrices were generated for a given fixed driver OD flow matrix.

A first analysis considers the driver OD flow matrix with random flows that was used in the previous sections. From this matrix, 200 passenger OD flow matrices were generated using the Dirichlet distribution.

For each drawn passenger matrix, we computed the Euclidean distance (see Appendix) between the driver OD flow matrix and the passenger OD flow matrix. Sensitivity analysis of optimal cost with respect to the Euclidean distance between the two matrices is examined. Optimal cost with respect to Euclidean distance for all matrices is shown in the north-west chart of Figure 7.

There is an increasing trend in the variation of the optimal social cost with the Euclidean distance. This is expected. Additionally, there is large variability within the results. For the same Euclidean distance, we have low and high optimal costs. If more passengers are assigned to an OD with higher travel cost, then optimal travel cost increases. If more passengers are assigned to an OD with lower travel cost, then optimal travel cost decreases. The results are robust with respect to the variation of the passenger OD flow distribution because the optimal cost variation is rather limited.

A second analysis considers a uniform driver OD flow matrix. In this case, the differences between the Dirichlet-generated passenger OD flow matrices and the driver OD flow matrix can be measured by either the Euclidean distance, or the Gini coefficient or the entropy.

The Gini coefficient is used to measure inequality in a distribution (see Appendix). A higher Gini coefficient indicates greater inequality. The coefficient ranges between 0 to 1, with 0 representing perfect equality (uniform distribution) and 1 representing perfect inequality. Entropy is used to measure a state of disorder, randomness, or uncertainty (see Appendix). The maximum entropy is obtained when the distribution is uniform and can be more than unity.

We have computed the variation in optimal cost with Euclidean distance, Gini coefficient and entropy for all 200 passenger OD matrices when the driver OD flows are assumed uniform. The three charts are in Figure 7. The trends are increasing when the Euclidean distance and the Gini coefficient are used. These results are expected. By contrast, when entropy is used, the trend is decreasing because the higher the entropy the closer is the passenger OD flow distribution to the uniform distribution. The variation in the optimal cost for the same Euclidean distance, Gini coefficient and entropy can be explained with the distribution of travel costs. Again, the optimal matching solution appears robust in terms of total cost with respect to variation of the passenger OD flow matrix on the driver OD flow matrix.

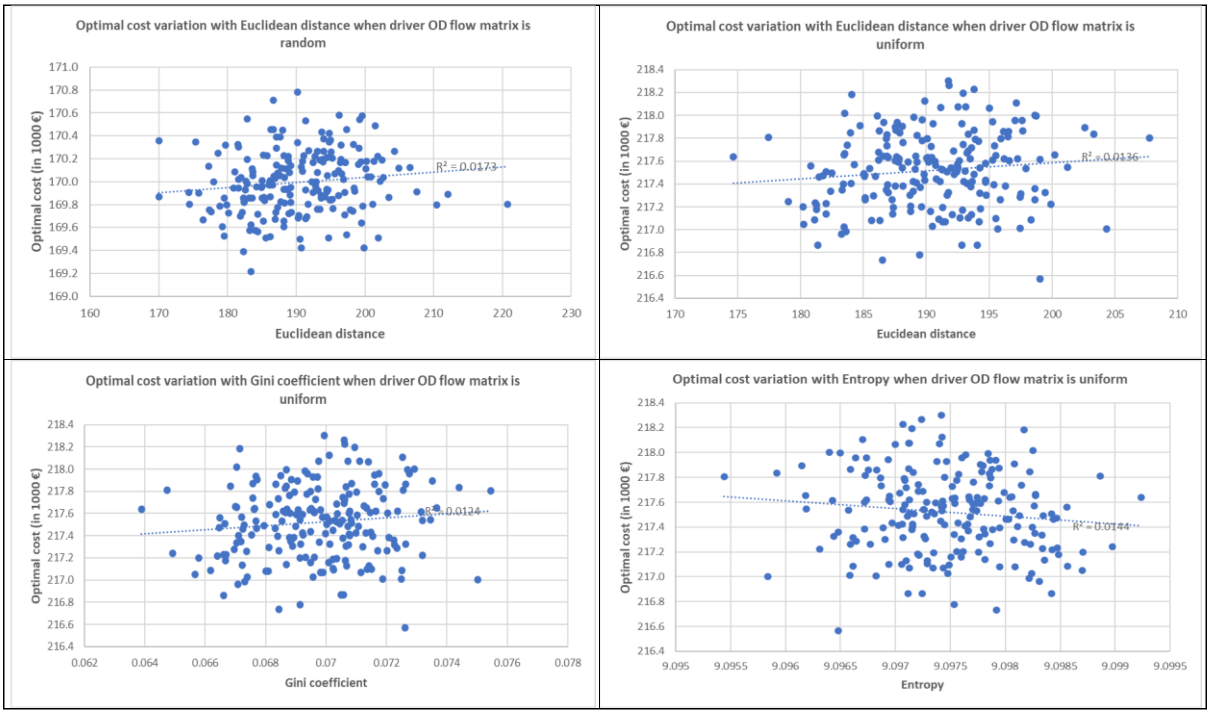


Fig. 7. Socially optimal matching with environmental costs: variation of optimal total social cost when passenger OD matrix diverges from driver OD flow matrix

5. Conclusions

In this paper, socially optimal matching and social planner's budget are investigated. Results from the triangle network and the Sioux Falls network help us understand the benefit of ridesharing in terms of savings in travel and emission costs. In the Sioux Falls network, with reference to the population of passengers and drivers who are willing to match, there is a total cost saving of around 18% on the all-travel-alone scenario, when in this scenario the share of those travelling by PT is 20%. Total vehicle-kilometres travelled by this population are reduced by around 35%. When matching is considered, there are two main drawbacks. The first is drivers and passengers who are required to travel alone, who account for 13% of total passengers and drivers willing to match. The second is detours, which account for 8% of total emissions.

The socially optimal matching solution where environmental costs are considered is better than the matching solution without considering environmental costs, since it offers a higher number of matchings and higher fuel and emission cost savings. The social planner's budget analysis shows that there is a net gain for the social planner. For each matching she has an average gain of 1.5 €, which can be used to operate the ridesharing scheme or can be re-invested elsewhere. The individual's preference of matching is not the same as the socially optimal matching, because individuals generally do not take into account the environmental costs when matching. Therefore, there is scope for a scheme like the one studied in this paper to reach socially optimal matching.

Our formulation and network analysis are restricted to the case where arc travel times are given exogenously and independent of the time of the day. When congestion is affected by ridesharing, it will be reduced by socially optimal matching, and the savings of travel costs (time and fuel) are expected to be higher. The endogenous congestion case needs to be formulated as one-step optimization problem with equilibrium constraints. The matching flows are the integer decision variables. However, for tractability reasons, they will be assumed as continuous. In a first stage, the problem can be tackled by iteratively solving the matching optimization problem and the traffic equilibrium problem with congestion. We have considered a uniform value of travel time. Future research may consider heterogeneity. Also pooling, where more than one passenger shares the ride with the driver, is left for future research.

For the implementation of the scheme, one needs to gather "lab" data on the value of time when travelling alone and when two or more individuals share the vehicle. Many parameters need to be estimated, including the psychological inconvenience and the risk attitude. For example, the rape and murder of a young passenger in the coastal city of Wenzhou led a partial boycott of Didi.

We expect the matching model to be used in many transport sectors, as it is already the case in many economic sectors, job, housing and marriage markets in particular. Taking account of global externalities in such markets may be worth economists' attention. We believe that our matching model is a necessary precursor of the matching model of autonomous vehicles. For now, the future of autonomous vehicles remains somewhat uncertain, but we believe transport economists' contribution is needed to set the stage. Note that the automobile industry and other stakeholders, like TNCs, do not necessarily have the same objective function (user equilibrium) as the transport economist (social optimum, including congestion and, possibly, environment). For example, it is well documented that congestion has increased in many cities after Uber has entered the market in the US (Schaller, 2021). The reason is that car occupancy is already low (for example 1.1 in Ile-de-France, Paris area) and it is likely to be even lower when the driver is a professional driver or a robot: an empty vehicle still creates congestion.

In the autonomous vehicle case, when the autonomous vehicles are individually owned, matching could work in a similar way as it does in this paper. Either the car picks up the passenger at the origin and does a minimal detour, or the passenger has to walk. The same discussion can be made at the destination. The setting differs when the vehicles are owned by a private or a government body (robot taxis). In this latter case the vehicle can circulate 24 hours. Of course, a new degree of freedom appears: where is each vehicle positioned for the next ride? We expect that in most cases the vehicles will be repositioned to meet the next ride. This latter issue requires a complex demand estimation and optimization problem, involving repositioning of vehicles from surplus areas to deficit areas, consideration of waiting times, and matching as the key step.

Acknowledgements

The authors would like to thank Pierre-André Chiappori, Jesse Jacobson, Alfred Galichon and Robin Lindsey for early discussions. We would like also to thank the participants to the seminars of CYU Cergy Paris Université and to the ITEA 2021 Annual Conference for their comments. The second author was supported by the NSFC-JPI UE joint research project "MAAT" (project no. 18356856).

Appendix

A.1 Dirichlet distribution

The Dirichlet distribution is a family of continuous multivariate probability distributions parameterized by a vector of positive reals (Ng et al., 2011). It is a multivariate generalisation of the Beta distribution.

Given the random vector $x = (x_1, \dots, x_n)$, where $x_i \epsilon (0,1), i = 1, \dots, n$, and $\sum_{i=1}^{n} x_i = 1$, the Dirichlet probability density function is:

$$f(x) = \frac{\Gamma(\sum_{i=1}^{n} \alpha_i)}{\prod_{i=1}^{n} \Gamma(\alpha_i)} \prod_{i=1}^{n} {x_i}^{\alpha_i - 1},$$

where $\alpha = (\alpha_1, \dots, \alpha_n)$ is a positive parameter vector and Γ is the Gamma function. The mean of the i-th component is:

$$\mathbb{E}[x_i] = \frac{\alpha_i}{\sum_{i=1}^{n} \alpha_i}.$$

When all $\alpha_i \to 0$, the distribution becomes noninformative. When $n = 2$, the Dirichlet distribution reduces to the Beta distribution.

A.2 Euclidean distance

The Euclidean distance provides a measure of the distance between two real-valued vectors (Horn and Johnson, 1990). Let $N_i^P = N_{Ai}^P + N_{Bi}^P$. The Euclidean distance between the driver and the passengers OD flow matrix is:

$$d = \sqrt{\sum_{i=1}^{M} \left(N_i^D - N_i^P\right)^2}.$$

A.3 Gini coefficient

The Gini coefficient provides a measure of inequality in a distribution. It is commonly used to measure inequality in a distribution of non-negative income or wealth. The coefficient ranges from 0 (or 0%) to 1 (or 100%), with 0 representing perfect equality and 1 representing perfect inequality.

Let $N_i^P = N_{Ai}^P + N_{Bi}^P$. The Gini coefficient related to the distribution of our passengers OD matrices can be computed using the formula (Dixon et al., 1988):

$$G = \frac{M+1}{M} - \frac{2\sum_{i=1}^{M}(M+1-i)N_i^P}{M\sum_{i=1}^{M} N_i^P},$$

where N_i^P, $i = 1, \dots, M$, are indexed in increasing order of OD flow.

A.4 Entropy

Entropy is used to measure a state of disorder, randomness, or uncertainty. Let $N_i^P = N_{Ai}^P + N_{Bi}^P$. The uncertainty in our generated passenger OD flow matrices is expressed by the following equation (Shannon, 1948):

$$H = -\sum_{i=1}^{M} p_i \log_2 p_i,$$

where H is the entropy and $p_i = N_i^P / \sum_{i=1}^{M} N_i^P$ is the probability of passengers being in OD pair i. For a given number of OD pairs M, H is maximum and equal to $\log_2 M$ when p_i are uniform and equal to $1/M$. This is also intuitively the most uncertain situation.

References

Beojone C.V., Geroliminis N. (2021) On the inefficiency of ride-sourcing services towards urban congestion. Transportation Research Part C: Emerging Technologies 124 102890, 1-20.

Campbell H. (2018) The Rideshare Guide. Everything You Need to Know About Driving for Uber, Lyft, and Other Ridesharing Companies. Skyhorse Publishing. New York NY.

Caulfield B. (2009) Estimating the environmental benefits of ride-sharing: A case study of Dublin. Transportation Research Part D: Transport and Environment 14 (7), 527–531.

Chen X., Di X. (2021) Ridesharing user equilibrium with nodal matching cost and its implications for congestion tolling and platform pricing. Transportation Research Part C: Emerging Technologies 129 103233, 1-29.

Dasgupta S., Papadimitriou C.H., Vazirani U.V. (2008) Algorithms. McGraw-Hill Higher Education, Boston MA.

de Palma A., Stokkink P., Geroliminis N. (2020) Influence of Dynamic Congestion on Carpooling Matching. THEMA Working Paper n°2020-12 CY Cergy Paris Université, France.

Diao M., Kong H., Zhao J. (2021) Impacts of transportation network companies on urban mobility. Nature Sustainability 4, 494-500.

Dixon P., Weiner J., Mitchell-Olds T., Woodley R. (1988) Erratum to 'Bootstrapping the Gini Coefficient of Inequality'. Ecology 69, 1307.

Dong Z., Leng M. (2021) Managing on-demand ridesharing operations: optimal pricing decisions for a ridesharing platform. International Journal of Production Economics 232.

Eiselt H.A., Sandblom C.-L. (2010) Linear Programming and its Applications. Springer-Verlag, Berlin.

European Commission (2019) Handbook on the External Costs of Transport. Version 2019 – 1.1. Directorate-General for Mobility and Transport, Brussels.

Gale D., Kuhn H., Tucker A.W. (1951) Linear programming and the theory of games. Chapter XII in: Koopmans T.C. (ed.) Activity Analysis of Production and Allocation. Wiley, New York.

Galichon A. (2018) Optimal Transport Methods in Economics. Princeton University Press, Princeton.

Horn R. A., Johnson C. R. (1990) Matrix Analysis. Cambridge University Press, Cambridge UK.

Korte B., Vygen J. (2008) Combinatorial Optimization. Theory and Algorithms. Springer-Verlag, Berlin.

LeBlanc L.J. (1975) An algorithm for the discrete network design problem. Transportation Science 9, 183-199.

Ng K.W., Tian G.-L., Tang M.-L. (2011) Dirichlet and Related Distributions: Theory, Methods and Applications. Wiley, Chichester UK.

Niculescu C.P., Persson L.-E. (2018) Convex Functions and Their Applications. A Contemporary Approach. Springer, Cham Switzerland.

Özkan E. (2020) Joint pricing and matching in ridesharing systems. European Journal of Operational Research 287(3), 1149-1160.

Roth A.E., Sotomayor M. (1990) Two-sided Matching. A Study in Game-Theoretic Modeling and Analysis. Cambridge University Press, Cambridge UK.

Schaller B. (2021) Can sharing a ride make for less traffic? Evidence from Uber and Lyft and implications for cities. Transport Policy 102, 1-10.

Shannon C. E. (1948) A mathematical theory of communication. The Bell System Technical Journal 27(3), 379-423.

Schrijver A. (1998) Theory of Linear and Integer Programming. John Wiley & Sons, Chichester.

Tikoudis I., Martinez L., Farrow K., Bouyssou C.G., Petrik O., Oueslati W. (2021) Ridesharing services and urban transport CO_2 emissions: simulation-based evidence from 247 cities. Transportation Research Part D: Transport and Environment 97 102923, 1-23.

Wang X., Wang J., Guo L., Liu W., Zhang X. (2021) A convex programming approach for ridesharing user equilibrium under fixed driver/rider demand. Transportation Research Part B: Methodological 149, 33-51.