

How to split the costs and charge the travellers sharing a ride? Aligning system’s optimum with users’ equilibrium

Andres Fielbaum^{1,*}, Rafał Kucharski², Oded Cats², and Javier Alonso-Mora¹

¹Department of Cognitive Robotics, Delft University of Technology, The Netherlands

²Department of Transport & Planning, Delft University of Technology, The Netherlands

*Corresponding author: a.s.fielbaumschnitzler@tudelft.nl

Abstract

Emerging on-demand sharing alternatives, in which one resource is utilised simultaneously by a circumstantial group of users, entail several challenges regarding how to coordinate such users. A very relevant case refers to how to form groups in a mobility system that offers shared rides, and how to split the costs within the travellers of a group. These are non-trivial tasks, as two objectives conflict: 1) minimising the total costs of the system, and 2) finding an equilibrium where each user is content with her assignment. Aligning both objectives is challenging, as users are not aware of the externalities induced to the rest.

In this paper, we propose protocols to share the costs within a ride so that optimal solutions can also constitute equilibria. To do this, we model the situation as a non-cooperative game, which can be seen as a game-version of the well-known *set cover problem*. We show that the traditional notions of equilibrium in game theory (Nash and Strong) are not useful here, and prove that determining whether a Strong Equilibrium exists is an NP-Complete problem, by reducing it to the *k-Exact-Cover* problem. Hence, we propose three alternative equilibrium notions (stronger than Nash and weaker than Strong), depending on how users can coordinate. For each of these equilibrium notions, we propose a (possibly overcharging) cost-sharing protocol that yields the optimal solutions equilibria.

Simulations for Amsterdam reveal that our protocols can achieve stable solutions that are close to the optimum, and that having a central coordinator can have a large impact.

Keywords: Transportation, Ridepooling, Cost-sharing, Price of stability, Set cover.

1 Introduction

Transport systems often face a tension between individual choices and global optimisation. As users need to share some limited resources, such as streets or vehicles, their decisions affect other travellers but such externalities are usually not internalised. There are some famous paradoxes illustrating this issue, such as the Braess paradox (Frank 1981), which predicts that building an extra road might make everybody worse through their selfish routing decisions, or the Down-Thomson paradox (Mogridge 1997), which states that new infrastructure for cars can also lead to an overall deterioration due to users switching from public transport to private modes. The problem of aligning users’ interests manifests itself in a novel way in on-demand shared systems, where requests are matched into groups, meaning that the route (thus waiting and in-vehicle times) depends on the circumstantial co-travellers. That is to say, if users can decide with whom to travel, they might induce externalities to the rest of the users, which implies that their individual interests might not be perfectly aligned with a global optimisation process.

Therefore, in such shared mobility systems it is crucial to distinguish between *optima* and *equilibria*. Optimal solutions are typically pursued by some central non-profit operator or authority, who wants the system to run as efficiently as possible, and refer to minimising the sum of all the costs involved. Equilibria (also called *stable* solutions) deal with users’ interests, namely to ensure that everybody is satisfied enough so that they will not coordinate somehow to change the solution. As we discuss below, there are different ways to define an equilibrium depending on how users are assumed to be able to coordinate.

The authority does have a tool so that users internalise (at least to some extent) the externalities they induce to the rest of the system: fares. In fact, when several users share a vehicle, an inevitable

question of how to split the monetary costs emerges, and there is no straightforward answer to it. Naive approaches might be to split the costs uniformly between the users, or proportionally to the distance between each user’s origin and destination. However, such ideas might lead to undesirable equilibria. If a user is travelling a long distance from her origin to her destination, then nobody would want to share the vehicle with her under a uniform division of the costs; the opposite situation would occur with fares that are proportional to the distance (a common scheme applied in ridepooling systems), as nobody would be willing to share a ride with someone requiring a short trip. Therefore, the question of which *cost-sharing protocol* should be implemented in order to align the users’ and the authority’s interests is far from trivial.

This paper is primarily devoted to addressing this research gap: How to formalise the problem of defining protocols to split the costs among users sharing a ride, and which protocols should be used to reach efficient solutions? To do so, we model the described situation as a formal game, in which each user can choose with whom to travel, as long as the selected co-travellers agree. We argue that the usual equilibrium notions (Nash and Strong Equilibria) do not capture appropriately how users can coordinate, which requires us to propose several alternative equilibria notions. For each of these notions, we propose a cost-sharing protocol (i.e., how to share the costs among the users within a group) that makes the optimal solution an equilibrium, so that the authority can propose the users how to match (optimally) in a way that they will be satisfied (Anshelevich et al. 2008). Finally, we test our ideas using real-life data in Amsterdam, with a batch of 400 travellers sharing rides.

The equilibrium notions we describe, as well as the corresponding cost-sharing protocols, might be utilised for any mobility system in which groups are formed on-demand and where a sufficient number of vehicles is always available (i.e., there are no rejected requests). The rest of the paper is written assuming a *ridepooling* system, i.e., a centralised service that matches travellers into groups and assigns vehicles to serve them. Ridepooling services are considered promising for the future of mobility, as they might keep many of the virtues that have made ride-hailing services popular, while reducing the increase in congestion that has been associated with those (Henao & Marshall 2019; Tirachini & Gomez-Lobo 2020; Diao et al. 2021; Wu & MacKenzie 2021). Aligning users’ and the system’s interests is crucial for this purpose, as one needs users to be interested in using the system, and the system to be able to effectively stimulate users to share so that congestion is indeed mitigated. Our findings also apply to *ridesharing* services, in which different riders coordinate to transport in a vehicle driven by one of them (Agatz et al. 2012; Chan & Shaheen 2012; Furuhata et al. 2013; Mourad et al. 2019; Enzi et al. 2021; Özkan 2020). For instance, Lu & Quadrifoglio 2019 study a similar model, arguing that when groups are matched, any of the riders within a group can be the driver. Although these systems have not been able to become large-scale yet, they might become a relevant part of future shared mobility systems, when coordination tools continue to evolve. As these systems might operate in non-centralised fashions, users’ choices become a structural component, so understanding their impact is crucial. Moreover, as we use generic cost functions, other examples of the so-called *sharing economy* could be modelled using the same framework (such as sharing a hotel room, a parking lot or a car, as exemplified by C. Chau & Elbassioni 2017).

The remaining of the paper is organised as follows. Section 2 reviews the most important previous works. Section 3 formalises the game and describes the equilibrium notions that we study. Section 4 proposes the respective cost-sharing protocols, which are tested numerically in section 5. Finally, section 6 concludes and proposes some lines for future research.

2 Related works

2.1 Cost-sharing protocols

The problem of how to split costs among different users that share one or more common resources, and value them differently, has been studied by researchers for decades. This is usually done in the context of game theory, i.e., where a number of players need to make some choices (such as which resources to demand), so that their decisions affect each other and are taken simultaneously. *Cost-sharing protocols* have been proposed for both cooperative and non-cooperative games. In the former, the aim is forming coalitions, defined as sets of users that decide to cooperate to form the best possible solution, so that the costs are split within each coalition; hopefully, there will be a single coalition formed by everybody. The classical cost-sharing models within the cooperative framework are reviewed by Moulin 2002, who also shows that this problem relates (via the resulting individual costs) to the question of how to split the resources, and by Jain & Mahdian 2007. In the non-cooperative version, each player decides individually, so that the cost-sharing protocol is meant to push the players towards equilibria that are

beneficial for everybody (some examples using this approach are the papers by Buchbinder et al. 2008; Arin et al. 2009; Ma et al. 2017). In the context of ridepooling (or ridesharing), this is a very relevant difference, as in the cooperative framework the cost of a user within a group does not depend only on her co-travellers, but also on the other members of the coalition, which might be inconvenient as the protocols can be difficult to understand (and thus to accept) by the passengers.

Most of the models are *anonymous*, meaning that the cost of a resource does not depend on which users are using it, but only on its total amount. This is the case for resource allocation problems, in which each user selects one or more resources (or an amount of them), and the total cost of each resource depends on the number of players (or the corresponding total amounts) selecting them (Harks & Miller 2011; von Falkenhausen & Harks 2013; Gkatzelis et al. 2016). A particularly relevant case is congestion games (Harks & V. Timmermans 2021), in which every user has to choose among some predefined set of resources (for instance, different routes connecting her origin and destination); cost-sharing protocols have been proposed for this setting by Gairing et al. 2020 and Gkatzelis et al. 2014.

Crucially, in on-demand transportation systems costs should not be anonymous, as the total distance driven by each vehicle depends directly on the origins and destination of its users. That is, the cost of the resources should depend not only on the number of users, but specifically on the set of users. This framework is analysed by Harks & von Falkenhausen 2014 in a general facility location problem, where each user has a specific set of available resources to choose, by Roughgarden & Sundararajan 2009, in a model where there is a single resource, and by Albers 2009, Roughgarden & Schrijvers 2016, and H. Chen et al. 2010 for network design problems where users choose which arcs they want. Scheduling problems are a particular type of non-anonymous model where cost-sharing protocols are needed: in such problems, jobs are to be assigned to available machines, so that each job-machine pair has an associated length, with the resulting cost of each machine depending on the length of the tasks assigned to it (von Falkenhausen & Harks 2011; Caragiannis et al. 2017).

The model by C. Chau & Elbassioni 2017 is more similar to ours. They too use non-anonymous costs, and differ from the papers referred in the previous paragraph as they consider a scheme in which resources are always available (they call this characteristic as considering *canonical resources* that are always replaceable). In the context of ridepooling, considering canonical resources means that we assume that the fleet is large enough to transport everybody, regardless of how users are grouped. As we discuss in section 3, such an assumption will be decisive when studying which passengers can deviate from their current group, as they will always have the chance to travel alone. The paper by C. Chau & Elbassioni 2017 differs from ours as they use cooperative game theory (which might be troublesome in the ridepooling context, as discussed above), and also as they take some usual cost-sharing protocols and study the worst possible equilibrium, instead of finding protocols that make the optimal solution an equilibrium.

As we explain in section 3, the game we study can be analysed as a game-theory version of the classical combinatorial problem *set cover*. A similar game has been studied by Escoffier et al. 2010, although they focus on cost-sharing protocols to minimize the cost of the worst possible equilibrium. Moreover, in their model the cost of a user within a group can increase if some co-travellers decide to join a different group, which is reasonable for the abstract model they study, but does not seem to be realistic for modelling a mobility system.

2.2 Sharing costs in on-demand transportation systems

Cooperative game theory has also been used for the specific problem of splitting the costs the users sharing a vehicle. Such an approach is followed by Lu & Quadrifoglio 2019, who study the case in which vehicles follow the shortest circuits to serve their passengers (i.e., solving a traveling-salesman-problem, TSP), which might not be the best case when users' costs are part of the decision. They focus on finding the so-called *nucleus* of the cooperative game, meaning that they minimise the maximal dissatisfaction among all the groups that are formed. Levinger et al. 2020 also study a similar framework as a cooperative game, and focus on how to compute the so-called *Shapley values*, which are known to be fair prices in such cooperative games, but are usually difficult to compute. Their main result is that when only vehicle-kilometers are taken into account and users are sorted a-priori, Shapley values can be calculated in polynomial time. Bistaffa et al. 2017 deals with both the optimal solution and the *kernel* of the cooperative game on a similar ridesharing system.

A different tool that has been widely used is *mechanisms design*, in which users offer prices as in an auction. Kleiner et al. 2011; Shen et al. 2016; Cheng et al. 2014; Bian et al. 2020 study mechanisms to match that are *incentive compatible*, i.e., in which each user's best strategy is to reveal their true interests: Kleiner et al. 2011 focus on a ridesharing system, Shen et al. 2016 on ridepooling, while

Cheng et al. 2014; Bian et al. 2020 on feeder-trunk systems where the on-demand component serves the last mile; in the latter, users do not share the vehicle simultaneously. Regarding non-shared ridehailing, Asghari et al. 2016 proposes bid mechanisms to assign drivers to riders, while J. Zhang et al. 2015 study a *discounted trade reduction mechanism* scheme to satisfy every agent.

Other studies use different techniques to set pricing that attain a stable matching among different agents: vehicles and users are matched by Rasulkhani & Chow 2019 in a generic many-to-one transport system (i.e., where many passengers can utilise the same route), while Peng et al. 2020; Hu et al. 2021 propose how to define the payments from passengers to drivers for ridesharing. L. Chen et al. 2018 consider a model in which users can choose among a set of options that offer different prices and pick-up times. Stable assignments between pairs of riders (or between one driver and one rider) are studied by Wang et al. 2018; H. Zhang & Zhao 2018; S. C.-K. Chau et al. 2020; Yan et al. 2020. Furuhata et al. 2015 study cost-sharing mechanisms for ridesharing, that are updated online as new passengers enter the system, guaranteeing that fares can never increase for a passenger. Ke et al. 2020 study the emerging market equilibria in both pooled and non-pooled systems, including drivers' and riders' interests.

In all, the relationship between efficient prices and stable/optimal assignments in on-demand mobility systems has been increasingly studied during last years. However, the pricing strategies have mostly relied either on prices that might depend on users travelling in other vehicles, or on auctions, so that the question about direct prices for a shared trip remains yet to be studied.

2.3 Game theory and flexible mobility systems

Flexible systems require deciding how to match vehicles and users, so that conflicts between the interests of different stakeholders usually emerge. Therefore, using game theory is a natural idea that has been followed by several papers in the past to study different issues related to these mobility systems. Hernández et al. 2018 consider an abstract model of carpooling, in which the users' decisions (or *strategies*) are whether to cooperate or not. If they do not cooperate, they can decide selfishly, but receive a punishment, whereas cooperating entails a reward. They focus on the evolution in time of users' decisions. Schroder et al. 2020 use game theory to understand the emergence of surge pricing in ridehailing systems, through a model in which drivers can decide when to turn-off their devices (a similar analysis is performed by Castillo et al. 2017, but without employing game theory). Kucharski et al. 2020 study the system-wide impact of users arriving late at their pick-up locations, which is also modelled using game theory, in which users' decide strategically how late to arrive, taken into account the annoyance of both waiting for other passengers and arriving late at the final destination. Jacob & Roet-Green 2021 studies a for-profit ridehailing system in which users strategically decide whether to travel solo, pooling, or not using the system at all.

In spite of its relevance, (non-cooperative) game theory has not been widely used to study the matching issues emerging in mobility systems in which users share the same vehicle.

2.4 Contribution

This paper's contribution is threefold: First, we identify that optimal ridepooling solutions are not always stable if travellers can choose with whom to share the vehicle, and formalise the problem as a non-cooperative game. Second, we show that the traditional notions of Nash and Strong Equilibria are not suitable for this problem, and propose three intermediate notions of equilibrium. The third and main contribution of this paper is proposing cost-sharing protocols, so that for each of the three equilibrium notions, optimal ways to group the users also constitute equilibria. The rules imposed by these protocols to split the costs within a shared trip depend only on the characteristics of the trip itself, and not on the other groups and trips in the system.

Additionally, we run numerical simulations for a real-life case from Amsterdam, that confirm our results, showing that optimal solutions are in fact stable for users. For each protocol, the best possible equilibrium yields total costs that are close to optimality regardless of the equilibrium notion, whereas that the price of anarchy can be significantly larger.

3 The game and the equilibria

3.1 The formal Co-Travellers Game CTG

Our problem is defined by a set of passengers $P = \{1, \dots, n\}$ and a set of feasible groups $\mathcal{G} = \{G_1, \dots, G_m\}$, each of them having a cost $c(G) \geq 0$. Groups $G \in \mathcal{G}$ are subsets of P , and all their subsets are assumed to be feasible as well, i.e. $G \in \mathcal{G}, H \subseteq G \Rightarrow H \in \mathcal{G}$. In particular, singletons are always feasible, i.e. $\{i\} \in \mathcal{G} \forall i \in P$.

Groups in \mathcal{G} are interpreted as sets of users that might travel in the same vehicle. The set \mathcal{G} is defined by exogenous conditions (such as the capacity of the vehicles, or declining groups that require too long detours), i.e., we assume that \mathcal{G} is fixed. We assume that there are enough vehicles to carry everybody.

Regarding costs, a natural assumption is that if $H \subseteq G$ then $c(G) \leq c(H)$ (and $c(\emptyset) = 0$). Differently from other related papers on the topic, we do not assume any type of supermodularity or submodularity, i.e., if G_1 and G_2 are disjoint groups and $G_1 \cup G_2$ is feasible, it might hold either that $c(G_1) + c(G_2) \leq c(G_1 \cup G_2)$ or that $c(G_1) + c(G_2) \geq c(G_1 \cup G_2)$, because the sign of this relationship represents how efficient it is to serve the whole group of passengers together (which depends, for instance, on how close are their origins and destinations).

We are interested in comparing different ways to match the users. A *matching* is a selection of groups $\{H_1, \dots, H_\eta\} \subseteq \mathcal{G}$, such that $\eta > 0$ is any integer and each $i \in P$ belongs to exactly one of these groups H . The optimal matching can be found using the ILP shown in Eq. (1), where $x_G = 1$ if and only if group G is selected to be executed, and the constraint ensures that each passenger is transported in exactly one group. Hence, we are looking for a mutually exclusive and collectively exhausting partitioning of the passengers population. Throughout the paper, we will refer to this ILP as Problem (1).

$$\begin{aligned} \min_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}} x_G c(G) & \tag{1} \\ \text{s.t.} \quad \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \end{aligned}$$

Remark: Mobility systems induce costs of different nature and to various agents. We opt to use an abstract representation $c(G)$ so that our model remains valid for any cost function. For instance, $c(G)$ can contain operators' costs, such as fuel and maintenance, and users' costs such as total travelling time or other specific ridepooling-related aspects (like unreliability, studied by Fielbaum & Alonso-Mora 2020; Alonso-González et al. 2020b, or the willingness to share the vehicle with strangers, studied by Alonso-González et al. 2020a; Ho et al. 2018, among others); one might even include in $c(G)$ other societal costs, such as congestion or emissions.

The discussion about stable solutions requires defining some individual costs. For each $G \in \mathcal{G}$ and $i \in G$, we denote by $c_i(G)$ the individual cost of i when she belongs in group G . We do not impose that $\sum_i c_i(G) = c(G)$, and we discuss when this is the case. These individual costs will be seen as exogenous when defining equilibrium notions in sections 3.2-3.3, and we will study how to define them in section 4. Deciding the individual costs is interpreted as defining the monetary fares for each user. The problem faced by an operator that is aiming for an optimal solution that is also stable can be seen as a two-level optimisation: First, the operator decides on some pricing scheme, and then the users choose how to match. **The question that needs to be solved by the system is then how to set prices so that users match in a way that yields low total costs.**

Problem (1) serves as the benchmark to compare how much worse is a resulting (stable) matching compared with the system optimum solution. Therefore, it is useful to see that Problem (1) can be read exactly as the set cover problem, which is known to be NP-Hard, and even further, that it is impossible to find a polynomial algorithm that provides a solution that approximates the optimal solution by a constant factor, unless $P = NP$ (Raz & Safra 1997). That is to say, our problem can be seen as a version of set cover in which each element of the set might decide which subset is covering it¹.

¹In the traditional set cover problem, elements of the universal set might be covered by more than one subset, which we do not allow here when we impose an equality in the constraint in Problem (1). However, it is straightforward to see that the problems are indeed equivalent, modifying the usual definition of set cover: For each $G \in \mathcal{G}$, for each $H \subseteq G$, if $H \notin \mathcal{G}$, we add H to \mathcal{G} with $c(H) = \min_{G' \in \mathcal{G}: H \subseteq G'} c(G')$. Doing so, there is always an optimal solution covering each element exactly once.

Although Problem (1) could be hard to solve (because set cover is NP-Hard), real-life ridepooling situations usually generate versions in which standard solvers are able to manage it (Alonso-Mora et al. 2017b; Kucharski & Cats 2020; Fielbaum et al. 2021a). Therefore, we shall assume in the rest of the paper that the optimal solution can be computed in reasonable time.

We can now formally define the underlying (non-cooperative) game, denoted as CTG: *Co-Travellers Game*. The players of CTG are the passengers p_1, \dots, p_n . Each passenger i can decide in which group she wants to travel, i.e., her possible strategies are the sets G in \mathcal{G}_i , defined as $\mathcal{G}_i = \{G \in \mathcal{G} : i \in G\}$.

Following the usual notation in game theory, we denote G_i the strategy (i.e., the group) chosen by i , and $G_{-i} = (G_j)_{j \neq i}$ the profile of strategies chosen by the other players. When i chooses G_i , there are two possibilities: either all the other players in G_i choose G_i as well, or at least one player does not. We assume that no one can be forced to join a group, so G is executed only in the first case, yielding a disutility function U :

$$U_i(G_i, G_{-i}) \begin{cases} c_i(G_i) & \text{if } G_j = G_i \forall j \in G_i \\ +\infty & \text{otherwise} \end{cases} \quad (2)$$

The $+\infty$ disutility in Eq. (2) represents the case in which the group is not executed, so i cannot travel.

We now discuss different equilibrium notions for this game, which are characterised by the level of coordination that is admitted among the users. For instance, the traditional Nash Equilibrium (NE) assumes that users cannot coordinate at all, so each user can only act unilaterally, whereas the traditional Strong Equilibrium (SE) assumes full coordination. We also propose three alternative notions that lay between those two². For any equilibrium notion, we are mostly interested in the *price of stability* (PoS), i.e., in the “best” possible equilibrium, meaning the equilibrium that yields lowest total costs, which is interpreted as the matching that the system should propose to the users. It is worth noting that the PoS is usually defined for NE (Anshelevich et al. 2008), but it can be computed for any equilibrium notion.

When we run numerical experiments, we will also look at the *price of anarchy* (PoA), which is the “worst” possible equilibrium, i.e., the one that yields the highest total costs, which represents the worst case scenario when users coordinate freely without any recommendation from the system.

3.2 Traditional notions of equilibrium

3.2.1 Traditional Nash Equilibrium (NE)

The most common notion in game theory is the Nash Equilibrium:

Definition 1. A profile of strategies $(G_i)_{i \in P}$ is a NE if no player can unilaterally improve her situation, i.e.

$$\forall i \in P, \forall G'_i \in \mathcal{G}_i, U_i(G_i, G_{-i}) \leq U_i(G'_i, G_{-i}) \quad (3)$$

In Definition 1 it is implicitly assumed that players are only opting for *pure strategies*. However, the usual notion of NE also admits *mixed strategies*, in which player i decides a probability distribution over \mathcal{G}_i in order to minimise the expected disutility. The following Lemma shows that in the CTG any NE is composed only by pure strategies because otherwise some players would face disutility $+\infty$ and could reduce it.

Lemma 3.1. Any Nash Equilibrium of the CTG is formed by pure strategies only.

Proof. Suppose player i chooses a mixed strategy, with $\text{Prob}(G_i = G) = p$ for some $G \in \mathcal{G}_i, p \in (0, 1), G \neq \{i\}$. Let $j \in G$ be a different player. Then

$$E(U_j(G)) \geq (1 - p) \cdot U_j(G | G_i \neq G) = (1 - p) \cdot +\infty = +\infty$$

That is to say, if j chooses G , then she faces an infinity expected disutility, because there is a chance that the group is not executed. Hence, G cannot be chosen by j with a non-zero probability in an equilibrium, because travelling alone yields a lower disutility. The same argument can now be applied to i . Given that we already know that j is assigning zero probability to G :

²This idea of proposing an ad-hoc equilibrium notion, between NE and SE, has also been proposed by Dosa 2018 for the so-called *Bin-packing problem*.

$$U_i(G_i, G_{-i}) = +\infty$$

Hence, i can improve her situation by assigning the mentioned probability p to $\{i\}$ instead of G , which proves that this situation is not an equilibrium. \square

We say that a profile of strategies $(G_i)_{i \in P}$ is *coordinated* if nobody chooses a group that is not being executed. Note that a coordinate profile of strategies is a matching (i.e., it satisfies the constraint in Eq. 1). Although non-coordinated situations are feasible, they shall not occur because any passenger would opt for travelling alone rather than choosing a group that is not being executed. This is formalised in the following Lemma for NE, and remain true for all the other notions of equilibrium we study below in this paper because they are stronger (that is, they also imply a NE):

Lemma 3.2. *If a profile of strategies is a Nash Equilibrium, then it is coordinated.*

Proof. Let $(G_i)_{i \in P}$ be a non-coordinated profile of strategies. Let i be a player choosing a group that is not being executed. Then i would be better-off by choosing $\{i\}$ instead, implying that $(G_i)_{i \in P}$ is not an equilibrium. \square

To simplify the notation, and because our analyses deal with users reaching equilibria, we drop the disutility U_i and will utilise the costs c_i directly, because they always coincide.

We now argue that a NE always exists. Let us consider a NE and a player i : Eq. (3) is trivially fulfilled by any G that is neither G_i nor $\{i\}$, because if i would switch to such a G , her group would not be executed and her cost would be $+\infty$. Therefore, Eq. (3) can be replaced by

$$\forall i \in P, c_i(G_i, G_{-i}) \leq c_i(\{i\}, G_{-i}) \quad (4)$$

Note that when player i travels by herself, her strategy is $G_i = \{i\}$ and Eq. (4) is fulfilled with an equality. Hence, everybody travelling alone constitutes a trivial NE for the CTG. We now define the set of groups in which Eq. (4) holds:

$$\mathcal{G}_{TNE} = \{G \in \mathcal{G} : \forall i \in G, c_i(G) \leq c_i(\{i\})\} \quad (5)$$

Any solution to Problem (1) that only selects groups from \mathcal{G}_{TNE} would be a NE, as groups from \mathcal{G}_{TNE} are defined such that Eq. (4) is fulfilled. Hence, we can find the PoS and the PoA by solving, respectively, problems (6) and (7):

$$\begin{aligned} \min_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}_{TNE}} x_G c(G) & \quad (6) \\ \text{s.t.} \quad \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \end{aligned}$$

$$\begin{aligned} \max_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}_{TNE}} x_G c(G) & \quad (7) \\ \text{s.t.} \quad \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \end{aligned}$$

Problem (6) results from constraining Problem (1) to consider only groups in \mathcal{G}_{TNE} , i.e. finding the best solution among this subset of groups, while Problem (7) switches from a minimisation to a maximisation problem, to find the worst solution.

The trivial NE mentioned above ($G_i = \{i\} \forall i \in P$) is troublesome, because it does not represent what is expected to happen in ridepooling services. If we think of such a service as provided by an app, a NE would imply that a user that is suggested to join a group can only choose to accept such a group or to dismiss it and travel alone. This occurs because NE is a weak notion, that forbids any type of coordination among users, which precludes accounting for the complexity in the equilibrium analysis that is introduced by the chance of sharing. All of this justifies analysing other types of equilibria.

3.2.2 Traditional Strong Equilibrium (SE)

An alternative usual notion in game theory, that represents the other extreme situation, is the Strong Equilibrium, in which each subset of players can coordinate:

Definition 2. *A matching $(G_i)_{i \in P}$ is a SE if no group of players can jointly improve their situation. As the utility of a player only depends on her co-travellers, this happens if and only if*

$$\forall G \in \mathcal{G}, \text{ either } \forall i \in G, G_i = G, \text{ or } \exists i \in G \text{ such that } c_i(G_i) \leq c_i(G) \quad (8)$$

Eq. (8) ensures that if a group G is not taking place, then it cannot happen that all the users from G would want to abandon their current groups to form G . This is a quite restrictive notion of an equilibrium because it permits any type of coordination between all the users in the game. Therefore, it does not come as a surprise that there are cases in which there is no such an equilibrium³. We show this in the following example, where we do not provide the explicit numerical costs of each group, but only how users rank them, as the comparison among groups is all that matters to determine whether a certain profile of strategies is an equilibrium:

Example 3.3. *Consider an instance of CTG with three players A, B and C . Travelling all of them together is the worst option for everybody. Everybody prefers travelling in pairs (regardless of the co-traveller) over travelling alone. Regarding pairwise trips, A prefers to travel with B over travelling with C , B prefers C over A , and C prefers A over B . Such an instance presents no SE: if they all travel alone, any pair would prefer to join, while if two players are travelling together, say A and B , then B would coordinate with C to create the group $\{B, C\}$, and analogous situations would occur for any other pair. Everybody travelling together is also not a SE because players would prefer to abandon the group.*

Not only an equilibrium might not exist. Determining whether that is the case is computationally intractable, unless all the groups are formed by one or two users.

Theorem 3.4. *a) Given an instance of CTG, determining whether a SE exists is NP-Complete. This is true even if groups' sizes cannot be larger than 3. b) On the other hand, if no group has more than 2 players, it is possible to determine the existence of a SE in polynomial time.*

Proof. We first argue that this problem is NP, by noting that it is possible to verify in polynomial time if a given profile of strategies is a SE. Indeed, we need to determine if Eq. (8) is fulfilled, which is done by checking once for each $G \in \mathcal{G}$ if the respective conditions described by Eq. (8) are met.

a) To prove that the problem in which groups can have 3 or fewer players is NP-Complete, we use a reduction from 3-Exact-Cover. The k -Exact-Cover is defined by a universe set S whose size is a multiple of k , and a collection of subsets $\mathcal{B} = b_1, \dots, b_Q$, with $|b_i| = k \forall i = 1, \dots, Q$. The question is if one can pick some of those subsets, such that they are all disjoint and cover the set S . The k -Exact-Cover is known to be NP-Complete for any $k \geq 3$ (Garey & Johnson 1979).

Given an instance of 3-Exact-Cover, we build an instance of CTG as follows:

- The set of players is $S^* = S \cup \{\alpha\}$, with α an additional artificial player.
- Each subset of S^* whose size is not larger than 3 is a feasible group.

In order to define the costs, please note first that, for this proof, we might drop the need that if $H \subseteq G$, then $c(H) \leq c(G)$. In fact, if our cost scheme does not meet this condition, we can add a large quantity D to $c_i(g)$ for each player i and for each group $g \in \mathcal{G}_i$. By doing so, the equilibria analysis remains unchanged because all individual costs are raised by the same amount, and the cost of each group g will have a part that is $D \cdot |g|$, such that if D is large enough then this proportional part will outweigh any other differences between groups of different size. Moreover, we do not need to define the costs numerically, as it suffices to show for each user how does she rank the different groups that she belongs to:

³When the system can decide on the prices, there is a trivial way to induce a SE, by sharing the total costs of a group uniformly among the users. After that, a greedy algorithm that repeatedly picks the group of non-selected users with the lowest average cost will output a SE. It is noteworthy that this mimics exactly the well-known polynomial greedy algorithm for set cover, implying that its cost is no larger than $\log(n+1)$ times the optimal (Chvatal 1979; Raz & Safra 1997). However, as discussed in the Introduction, such a way to split the costs would probably not be accepted by users travelling short distances.

- Groups of size 3 corresponding to some b_i (i.e., that comes from the instance of the original 3-cover problem) are equally ranked as the preferred groups for all their members.
- Groups of size 1 are the less convenient ones for everybody.

To rank the remaining groups, i.e., groups of size 3 that do not correspond to any b_i , and pairs, we first introduce a notion of individual preferences for each user: We say that player i prefers player j_1 over j_2 if and only if

$$j_1 - i \pmod{n} < j_2 - i \pmod{n} \quad (9)$$

And we denote this situation by $j_1 \prec_i j_2$. Eq. (9) can be interpreted as if all users were displayed in a circle, such that i always prefers the first player that appears to her right. Therefore, $i + 1 \pmod{n}$ is the most preferred and $i - 1 \pmod{n}$ is the least preferred player for player i .

For player i , the remaining groups are sorted *lexicographically* according to the relationship \prec_i . That is to say, consider two different groups $J = (i, j_1, j_2)$ and $K = (i, k_1, k_2)$, such that $j_1 \prec_i j_2$ and $k_1 \prec_i k_2$. Then, i prefers J over K if and only if

$$j_1 \prec_i k_1 \text{ or } (j_1 = k_1 \text{ and } j_2 \prec_i k_2) \quad (10)$$

Intuitively, for player i the most relevant aspect to evaluate the group that she belongs to is the closest of her co-players. If two groups present the same closest co-player, then the remaining co-player is relevant, following the same order. Such a first criterion is also the decisive one when there is a group of two users involved, i.e., if the same group J is compared with the group $K' = (i, k)$, i prefers J if and only if

$$j_1 \prec_i k \text{ or } j_1 = k \quad (11)$$

Finally, if i has to choose between (i, j) and (i, k) , she will chose (i, j) if and only if $j \prec_i k$. The details that prove that this reduction works can be found in the appendix.

b) To prove that the problem of determining whether a SE exists is polynomial when all the groups have 1 or 2 users, we show that such a problem is equivalent to the well-known *stable-roommate problem*, which can be solved in polynomial time (Gusfield 1988). The details proving such an equivalence can be found in the appendix. \square

It is noteworthy that determining whether a general game has a SE can be harder than being NP-Complete (Gottlob et al. 2005), because verifying that a given profile is a SE is not as simple as here. The main difference is that in CTG we discard all the non-coordinated profiles of strategies.

The reason why such a notion of equilibrium is not appropriate for this game is that it cannot take place in ridepooling nor ridesharing systems. If hundreds or thousands of users are involved, it is impossible that all of them are able to coordinate just to perform a trip. That is, involving every possible coordination to define stability is too strong.

3.3 Alternative notions of equilibrium

The two most traditional notions of equilibrium are not suitable for this model. NE is too weak, and SE is too strong. Therefore, we now propose three alternative definitions for equilibrium in the ridepooling context. All of them lie in-between the two previous ones, permitting some level of joint decisions among the users. When we denote them with abbreviations, we will use first a letter “R” to signal that they represent ridepooling situations.

3.3.1 Ridepooling Hermetic Equilibrium (RHE)

The first alternative notion of equilibrium recognises that a user could coordinate with those other users that are already related to her: the ones with whom she is currently sharing.

Consider $G \in \mathcal{G}$ and $H \subsetneq G$. We say that H wants to leave G if $\forall i \in H, c_i(H) < c_i(G)$, and G is *hermetic* if no such an H exists. Formally, G is hermetic if and only if:

$$\forall H \subsetneq G, \exists i \in H \text{ such that } c_i(G) \leq c_i(H) \quad (12)$$

This idea permits coordination between users, but only when they are in the same group. Verifying if a group is hermetic can be done by testing for all its proper subsets if they want to leave or not.

Definition 3. A matching $(G_i)_{i \in P}$ is a hermetic ridepooling equilibrium if every G_i is hermetic.

It is evident that having each traveller by herself is a RHE. However, it is not necessarily unique. To find the PoS and PoA, we constraint Problem (1) in a similar way as for NE, but with tighter restrictions. Let us define \mathcal{G}_H as the set of hermetic groups, which is found just by pruning the non-hermetic groups. The PoS and PoA are found solving problems (13)-(14), that are analogous to (6)-(7), constraining the set of groups to \mathcal{G}_H .

$$\begin{aligned} \min_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}_H} x_G c(G) & \quad (13) \\ \text{s.t.} \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \end{aligned}$$

$$\begin{aligned} \max_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}_H} x_G c(G) & \quad (14) \\ \text{s.t.} \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \end{aligned}$$

3.3.2 Ridepooling Unmergeable Equilibrium (RUE)

We now propose an alternative notion of equilibrium that admits some trivial Pareto improvements: merging groups when it is better for everybody. Allowing for these simple movements prevents to have as an equilibrium the matching in which everybody travels alone.

We say that two disjoint groups G_1, G_2 are *mergeable* if $G_1 \cup G_2 \in \mathcal{G}$ and

$$\forall i \in G_1 \cup G_2, c_i(G_1 \cup G_2) \leq c_i(G_i), \text{ and } \exists i \in G_1 \cup G_2 \text{ such that } c_i(G_1 \cup G_2) < c_i(G_i) \quad (15)$$

This represents a way in which users from two different vehicles can coordinate, but only if everyone benefits from doing so. As this is a Pareto improvement (i.e. some players get better-off while nobody loses), these mergers might be suggested for an external controller.

Definition 4. A matching $(G_i)_{i \in P}$ is a ridepooling unmergeable equilibrium if it is a NE and no two different groups G_i, G_j are mergeable.

Such an equilibrium always exists, as shown in Theorem 3.5.

Theorem 3.5. Any instance of CTG admits a RUE.

Proof. In the appendix. □

As this equilibrium condition does not depend on single groups but rather on the chance of merging them, we cannot just restrict \mathcal{G} to find PoS and PoA. Instead we first identify all the pairs of groups that are mergeable, and then use an extra constraint to prevent that they both occur simultaneously. In Algorithm 1 we build the set \mathcal{M} of mergeable groups, i.e., each element of \mathcal{M} is a duple formed by two groups that are mergeable.

Algorithm 1: Construction of \mathcal{M} .

```

 $\mathcal{M} = \emptyset;$ 
for all  $G_1, G_2 \in \mathcal{G}_{TNE}$  do
  if  $G_1 \cap G_2 = \emptyset$  and  $G_1, G_2$  are mergeable then
     $\mathcal{M} \leftarrow \mathcal{M} \cup (G_1, G_2);$ 
  end if
end for
Output  $\mathcal{M};$ 

```

If $(G_1, G_2) \in \mathcal{M}$, both occurring simultaneously would mean that the matching is not a RUE. Therefore, we prevent that from happening in the ILP by adding extra constraints ensuring that the respective binary variables cannot both take a value of 1. This is shown in the inequalities represented by the last lines in Eqs. (16)-(17), that find the PoS and PoA for this type of equilibrium.

$$\begin{aligned}
& \min_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}_{\mathcal{T}N\mathcal{E}}} x_G c(G) & (16) \\
& \text{s.t.} \quad \sum_{G:i \in G} x_G = 1 & \forall i \in P \\
& \quad \quad \quad x_{G_1} + x_{G_2} \leq 1 & \forall (G_1, G_2) \in \mathcal{M}
\end{aligned}$$

$$\begin{aligned}
& \max_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}_{\mathcal{T}N\mathcal{E}}} x_G c(G) & (17) \\
& \text{s.t.} \quad \sum_{G:i \in G} x_G = 1 & \forall i \in P \\
& \quad \quad \quad x_{G_1} + x_{G_2} \leq 1 & \forall (G_1, G_2) \in \mathcal{M}
\end{aligned}$$

3.3.3 Ridepooling Semi-individual Equilibrium(RSIE)

One of the reasons why NE recognises everyone travelling alone as an equilibrium, is that individual (unilateral) movements are very restrictive in the formal definition of CTG. A natural extension for the idea of individual movements is that a single player switches group, with the players from the receiving group having the chance of accepting her if everybody improves or stays the same. With this idea in mind, we can provide a formal definition of an equilibrium that is not fully, yet nearly, individual (because of the users that have to accept the moving player):

Let $G_1, G_2 \in \mathcal{G}$ and disjoint. We say that G_1, G_2 are *individually unstable* if $\exists i \in G_1$ such that $G_2 \cup \{i\} \in \mathcal{G}$ and the two following conditions are fulfilled:

1. $c_i(G_2 \cup \{i\}) < c_i(G_1)$
2. $\forall j \in G_2, c_j(G_2 \cup \{i\}) \leq c_j(G_2)$

The first condition entails that i wants to move from G_1 to G_2 , while the second condition ensures that everyone in G_2 is willing to accept i . This equilibrium can be seen as the one expected when the service is offered through an app, such that each user can take only two actions: i) leaving her current group (as for NE), and ii) joining a new group; the latter shall take place only if all users of the mentioned new group accept her.

Definition 5. A matching $(G_i)_{i \in P}$ is a ridepooling semi-individual equilibrium if no G_1, G_2 are individually unstable.

The same counter-example 3.3 that revealed that there are instances of CTG with no SE is also a scheme with no RSIE.

Theorem 3.6. Given an instance of CTG, determining whether a RSIE exists is NP-Complete. This is true even if groups' sizes cannot be larger than 3.

Proof. The same arguments used to prove Theorem 3.4 a) are valid for this theorem, as it involved only the kind of changes admitted by this equilibrium notion. \square

Similar to RUE, finding the PoS and PoA under this notion of equilibrium requires adding some constraints to the ILP. These constraints might lead to an empty set, if no RSIE exists. We first need to identify the pairs of sets that are individually unstable, which are kept in the set \mathcal{S} built in Algorithm 2. Here we assume explicitly that $\emptyset \in \mathcal{G}$, entailing that the RSIE is also a NE, when taking $G_2 = \emptyset$ in Algorithm 2.

Algorithm 2: Construction of \mathcal{S} .

```

 $\mathcal{S} = \emptyset;$ 
for all  $G_1, G_2 \in \mathcal{G}$  do
  if  $G_1 \cap G_2 = \emptyset$  and  $G_1, G_2$  are individually unstable then
     $\mathcal{S} \leftarrow \mathcal{S} \cup (G_1, G_2);$ 
  end if
end for
Output  $\mathcal{S};$ 

```

Once the set \mathcal{S} has been computed, the PoS and PoA for this type of equilibrium are obtained solving problems (18)-(19).

$$\begin{aligned}
 \min_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}} x_G c(G) & \tag{18} \\
 \text{s.t. } \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \\
 x_{G_1} + x_{G_2} \leq 1 & \quad \forall (G_1, G_2) \in \mathcal{S}
 \end{aligned}$$

$$\begin{aligned}
 \max_{x_G \in \{0,1\}} \sum_{G \in \mathcal{G}} x_G c(G) & \tag{19} \\
 \text{s.t. } \sum_{G: i \in G} x_G = 1 & \quad \forall i \in P \\
 x_{G_1} + x_{G_2} \leq 1 & \quad \forall (G_1, G_2) \in \mathcal{S}
 \end{aligned}$$

3.4 Synthesis and positioning of the different notions of equilibria

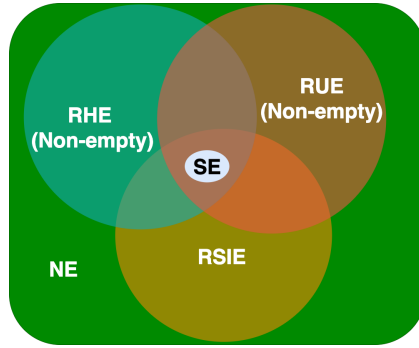


Figure 1: Venn diagram of the five notions of equilibria. NE, RHE and RUE are never empty, whereas RSIE and SE might be empty.

We have shown that NE is too weak as an equilibrium notion (i.e., it is too easy to fulfill) and that SE is too strong. We have therefore proposed three extra definitions for equilibria that lie in-between, depicted in Figure 1 together with NE and SE. These three alternative equilibria represent different types of coordination among users: RHE admits full coordination between users within the same group, RUE admits a simple type of Pareto improvement, and RSIE accepts individuals moving from one group to another, if the users from the latter are willing to accept her.

There is always at least one RHE and one RUE, entailing that the set of NE cannot be empty. The SE and RSIE sets, on the other hand, might be empty and it is an NP-Complete problem to determine if that is the case.

None of the intermediate notions of equilibrium is contained in the others. Consider the example 3.3: everybody travelling alone is a RHE, but it is not RUE nor RSIE (because travelling in pairs is better); the three users travelling together is a RUE (because there is only one group, so no possible mergers) but it is not RHE nor RSIE (because any player could opt to leave the group and travel alone). An example of a RSIE that is no RHE nor RUE is provided in the Appendix.

It is noteworthy that the intersections among the groups create new types of equilibria, in which two (or three) of the equilibria conditions are held simultaneously. For the sake of simplicity, we are not studying those intersections explicitly. Note that any of those intersection requiring RSIE might be empty. Moreover, there might also be no matching that satisfies both RHE and RUE: an example (with five players) for this is provided in the Appendix; however, one of the cost-sharing protocols we study in following section does make optimal solutions both a RHE and RUE.

4 Cost-sharing protocols to induce good equilibria

So far we have assumed that $c_i(G)$ is exogenous to the problem. However, from a policy point of view, one is interested not only in *predicting* which equilibria might emerge, but also in *steering towards* the best possible equilibria. Therefore, the controller of the system can determine the individual costs $c_i(G)$ for each user i and each group $G \in \mathcal{G}_i$ (for instance, via the monetary fares), in order to induce efficient outcomes.

From now on, we assume that only $c(G)$ (the total cost of the group) is given for each G . We aim to find a so-called *cost-sharing protocol*, i.e., to define how to split the costs amongst the players in G by defining the respective individual costs $c_i(G)$. Following Christodoulou et al. 2020, we will say that such protocols are:

- **Budget balanced** if they cover exactly the cost of the group, i.e.

$$\forall G \in \mathcal{G}, \sum_i c_i(G) = c(G) \quad (20)$$

- **Overcharging** if they are equal or in excess of the cost of the group, i.e.

$$\forall G \in \mathcal{G}, \sum_i c_i(G) \geq c(G) \quad (21)$$

Note that here we are not dealing with the profit of the system. The idea of an overcharging protocol is that the set of individual costs can effectively push users to decide on groups that are efficient from a global point of view.

As we aim at proposing pricing schemes that are understandable by the users, we will consider only *oblivious* protocols (as defined by Christodoulou et al. 2017), meaning that $c_i(G)$ depends only on i and G , and not on the other feasible and selected groups in \mathcal{G} . To be precise, an oblivious⁴ protocol in this context defines $c_i(G)$ as a function of the vector $(c(H))_{H \subseteq G}$. Moreover, for the different protocols we propose, all such functions can be computed in polynomial time.

In this section, for each of proposed equilibria (RUE, RHE or RSIE) we propose a cost-sharing protocol that make any optimal matching an equilibrium (i.e., PoS= 1). Before introducing the different protocols, it is useful to show and prove the following intuitive Lemma.

Lemma 4.1. *Let $\{c_i(G) : i \in P, G \in \mathcal{G}\}$ be a budget-balanced cost-sharing protocol. Then any optimal matching does not contain pairs of mergeable groups.*

Proof. Consider G_1 and G_2 part of an optimal matching. If G_1 and G_2 were mergeable (Eq. 15), then the total cost of $G_1 \cup G_2$ would be strictly lower than $c(G_1) + c(G_2)$, which contradicts the optimality of the matching. \square

We now proceed to introduce the different protocols.

4.1 Externality-based protocol

There is a vast literature that studies how incorporating externalities into pricing might induce optimal equilibria. The so-called Pigouvian taxes (Cremer et al. 1998) for markets regulation, the VCG-mechanisms for auction designs (Nisan & Ronen 2007), or the incremental cost-sharing protocols for demand games (Moulin 2008) are some of the most relevant examples. In CTG, however, doing so presents some difficulties:

- The externalities induced to the whole system depend on the complete set of passengers and groups, which would require violating the principle of proposing only oblivious protocols.
- We have defined several notions of equilibria, so making an optimal matching a NE, instead of a stronger notion, is not good enough.

⁴As discussed in Section 2, using oblivious protocols precludes us from using a cooperative game theory approach. Moreover, requiring the protocols to be oblivious also precludes using trivial ones to reach PoS=1. For instance, one could set null individual costs when users select the group corresponding to the optimal solution, and a positive cost otherwise. However, this would not be oblivious, as the costs $c_i(G)$ would depend on which is the optimal solution, an information that does not depend solely on G .

Notwithstanding the above, we now show how to define costs according to some of the externalities (namely those induced to the members of the same group), using an oblivious protocol and making any optimal matching a RSIE. However, such a protocol will not be budget balanced.

Definition 6. *Let G be a group. The **externality-based** protocol charges to each i in G the extra cost that its inclusion induces to the group, i.e.*

$$c_i(G) = c(G) - c(G \setminus \{i\}) \quad (22)$$

Theorem 4.2. *If the externality-based protocol is used, then every optimal matching is a RSIE.*

Proof. Let $(G_i)_{i \in P}$ be an optimal set of groups, and take player i and group $G_j \neq G_i$ such that $G_j \cup \{i\} \in \mathcal{G}$. We will prove that i does not want to move to G_j . If i moves to G_j , her new costs would be:

$$c_i(G_j \cup \{i\}) = c(G_j \cup \{i\}) - c(G_j) \quad (23)$$

The difference in costs for i is found by comparing Eqs. (22) and (23):

$$c(G_j \cup \{i\}) - c(G_j) - [c(G_i) - c(G_i \setminus \{i\})] = c(G_j \cup \{i\}) + c(G_i \setminus \{i\}) - [c(G_j) + c(G_i)] \quad (24)$$

The last expression corresponds exactly to the total changes in the global costs of the system, which are positive because $(G_i)_{i \in P}$ is optimal. Hence, player i would face positive extra costs if moving to any other group. □

Clearly, this protocol is not budget-balanced. In the presence of economies of scale, represented by sub-modular functions (i.e., in which adding an extra passenger is cheaper when the group is larger), the marginal individual costs $c_i(G)$ might become very low for large groups G , thus being far from sufficient to cover the real group costs. Note that this low individual costs represent effectively that larger groups should be prioritised when there are economies of scale. Thus, the system would require subsidisation, which is the usual case whenever there are scale economies in transport systems (Jara-Díaz 2007). If it is not possible to have subsidies, the problem is partially solved with Corollary 4.2.1, that shows that the externality-based split protocol can be adapted to be overcharging. However, it is no longer oblivious, but *resource-aware*, meaning that the costs within a group G might depend on all the possible costs $c(H)$ for $H \in \mathcal{G}$ (regardless of the actual groups that are chosen by passengers not in G). In practice, the only not-oblivious information that is used to make the protocol overcharging is $\max_{H \in \mathcal{G}} c(H)$.

Corollary 4.2.1. *There is an overcharging resource-aware protocol, that makes every optimal matching a RSIE. The only not-oblivious information required is $\max_{H \in \mathcal{G}} c(H)$.*

Proof. Let $D \geq \max_{H \in \mathcal{G}} c(H)$. Define $c_i(G)$ as if it was externality-based, but adding D , i.e.

$$c_i(G) = c(G) - c(G \setminus \{i\}) + D \quad (25)$$

As D is large enough, then individual costs are enough to cover the real ones. And because it is a fixed cost (every player i has to pay it regardless of the group), the analyses regarding equilibria from Theorem 4.2 are not affected. □

Remark: Theorem 3.6 states that determining whether a RSIE exists is NP-Complete, and we have just proved that if the externality-based protocol is used, not only does a RSIE exist, but actually every optimal solution is a RSIE. These two facts are not in contradiction, as the instances of the NP-Complete problem are defined by not only the group costs $c(G)$ but also exogenous individual costs $c_i(G)$. On the other hand, in Theorem 4.2 we use a specific definition for the individual costs, so that the instance is only defined by the group costs.

4.2 Residual-based protocol

We now study an idea that might look more natural, in which the costs faced by a user are directly related to the cost of the respective private trip. For this, we define the *residual cost* of a group $\Delta c(G)$ as the difference between the individual and the group costs:

$$\Delta c(G) = c(G) - \sum_{i \in G} c_i(\{i\}) \quad (26)$$

Such residual costs can be positive or negative, depending on whether it is efficient to group those users together. In general, one might expect that they are negative for the groups that do take place, as otherwise it is not reasonable to form that group. Nevertheless, we do not assume that in what follows.

Definition 7. *Let G be a group. A **residual-based protocol** consists of sharing only the residual costs, i.e., there exists residual prices $p(i, G) \forall i \in G$ such that $c_i(G) = c_i(\{i\}) + p(i, G)$. All residual prices $p(i, G)$ have the same sign and must fulfil:*

$$\sum_{i \in G} p(i, G) = \Delta c(G) \quad (27)$$

Note that requiring every $p(i, G)$ to have the same sign and to fulfil Eq. (27) implies that

$$\forall i \in G, \text{sign}(p(i, G)) = \text{sign}(\Delta c(G)) \quad (28)$$

There might be several different residual split protocols, depending on the definition of $p(i, G)$. Some simple ideas might include making $p(i, G)$ proportional to $c_i(\{i\})$, or just a uniform division. In any of which, users first pay a cost that depends solely on their particular travel characteristics, and only the residual component depends on the rest of the group. These protocols are obviously budget-balanced, and we now show that any optimal matching is a RUE when they are applied.

Theorem 4.3. *Assume that any residual split protocol is applied. Then, every optimal matching is a RUE.*

Proof. Let $(G_i)_{i \in P}$ be an optimal matching. We need to prove that it is a NE and that no two different groups are mergeable.

To see that it is indeed a NE, consider some G_i with more than one passenger (groups of size 1 do not need to be analysed). Because this is an optimal matching, the cost of $c(G_i)$ cannot be larger than all of its members travelling alone, i.e.

$$c(G_i) \leq \sum_{j \in G_i} c_j(\{j\}) \quad (29)$$

Eq. (29) entails directly that $\Delta c(G_i) \leq 0$ (due to Eq. 26), which implies that $p(j, G_i) \leq 0 \forall j \in G_i$ (Eq. 28), and thus $c_j(G_i) \leq c_j(\{j\})$. That is, every player j prefers her current situation rather than travelling alone, which is the definition of a NE.

The proof that no two groups can be mergeable follows directly from Lemma 4.1. □

For the numerical simulations in section 5, we will use residual prices that are proportional to the individual costs, i.e.

$$p(i, G) = \Delta c(G) \cdot \frac{c\{i\}}{\sum_{j \in G} c_j(\{j\})} \quad (30)$$

Which implies that total costs for each user will be proportional to their private costs.

4.3 Subgroup-based protocol

A different way to combine individual and group costs is based on looking at the subgroups each player might belong to. Consider a group G and a subgroup $H \subseteq G$. If H presents a low average cost (in comparison with G), then players that belong in H should not be charged much, as they are not directly inducing the costs on G . On the other hand, consider a player i such that every subset $H \subsetneq G$ containing i presents a higher average cost than G : in that case, player i would be directly benefited

when G takes place, so she should pay more than the average. With this idea in mind, let us propose algorithm⁵ 3 that takes as input the group G and outputs, for each player i , a subgroup-based cost $z_i(G)$ and an associated subgroup $\varphi_i(G)$.

Algorithm 3: Determining subgroup-based costs and the associated subgroups.

Input: G
 $W = G$; % W contains the users whose costs have not been determined yet
while $W \neq \emptyset$ **do**
 $J = \operatorname{argmin}_{H' \subseteq W} \frac{c(H')}{|H'|}$;
 for all $i \in J$ **do**
 $z_i(G) \leftarrow \frac{c(J)}{|J|}$, $\varphi_i(G) \leftarrow J$;
 end for
 $W \leftarrow W \setminus J$;
end while
Output: $z_i(G), \varphi_i(G), \forall i \in G$

The associated subgroups $\varphi_i(G)$ represent which group is defining the cost of each i . It is relevant to remark a straightforward property fulfilled by these subgroups: they form a partition of G , that is to say, two conditions are met:

1. $\forall i, j \in G$, either $\varphi_i(G) \cap \varphi_j(G) = \emptyset$ or $\varphi_i(G) = \varphi_j(G)$.
2. $\bigcup_{i \in G} \varphi_i(G) = G$.

Utilising costs z_i would not necessarily induce a budget-balanced protocol. Let us define the excess $e(G)$ as:

$$e(G) = c(G) - \sum_{i \in G} z_i(G) \quad (31)$$

This excess might be positive or negative. To define the actual protocol, we modify the payoffs z_i to achieve the budget-balanced property. If the excess is positive, we split the remainder uniformly among the users of the group (for our results below regarding optimal matching being equilibria, this remainder could be split in any way):

$$c_i(G) = z_i(G) + \frac{e(G)}{|G|} \text{ if } e(G) > 0 \quad (32)$$

When $e(G) < 0$, we sort the $\varphi_i(G)$ according to their average costs (from the cheapest to the most expensive), resulting in $\varphi^1, \dots, \varphi^Q$, with Q the number of distinct subgroups $\varphi_i(G)$. Let us denote the respective costs $z^i = c(\varphi^i)/|\varphi^i|$ (we are omitting the explicit reference to G to ease the notation). Define I_1 as the smallest index such that $z^{I_1} \geq c(G)/|G|$ (such an index exists because the excess is negative). We then diminish the costs of φ^j for $j \geq I_1$ to be equal to the average cost of G until we reach zero excess. The index I_2 representing the last φ^j whose cost is adjusted to $c(G)/|G|$ can be characterised as the largest index such that

$$\sum_{i=1}^{I_1-1} z^i |\varphi^i| + \sum_{i=I_1}^{I_2} \frac{c(G)}{|G|} |\varphi^i| + \sum_{i=I_2+1}^Q z^i |\varphi^i| \geq c(G) \quad (33)$$

The definition of I_2 through Eq. (33) implies that if we set the costs for users in φ^{I_2+1} to be equal to $c(G)/|G|$, then the individual costs would not be enough to cover the group's costs. However, such costs might still be reduced from the original z^{I_2+1} , till the zero excess is reached. The precise expression for the budget-balanced costs in the case of negative excess is defined by:

$$\forall k \in \varphi^i, c_k(G) = \begin{cases} z^i & \text{if } i < I_1 \text{ or } i > I_2 + 1 \\ \frac{c(G)}{|G|} & \text{if } i \in \{I_1, \dots, I_2\} \\ \frac{c(G) - (\sum_{j=1}^{I_1-1} z^j |\varphi^j| + \sum_{j=I_2}^{I_2} \frac{c(G)}{|G|} |\varphi^j| + \sum_{j=I_2+2}^Q z^j |\varphi^j|)}{|\varphi^i|} & \text{if } i = I_2 + 1 \end{cases} \quad (34)$$

⁵We choose this writing of the algorithm because it eases the understanding. The algorithm can be adapted to run in polynomial time: Sort in a list all groups in \mathcal{G} according to their average costs. Then, for each G , go through the said list, selecting the groups that are subsets of G and that only contain elements that have not been assigned yet.

The last case in Eq. (34) represents how to diminish the costs for users in φ^{I_2+1} to make the excess equal to zero, which is not enough for $c_k(G)$ to reach $c(G)/|G|$ (by definition of I_2). Note that, in this case of negative excess, costs can only be diminished, which implies that

$$c_i(G) \leq z_i(G) \forall i \in P \quad (35)$$

Definition 8. Let G be a group. The **subgroup-based protocol** charges to each i in G the costs according to Eq. (32) if $e(G) \geq 0$, or to Eq. (34) if $e(G) \leq 0$.

By construction, this protocol is budget balanced. The following theorem states that using it reaches a price of stability equal to 1, if the system is commanded by the RHE or RUE notions (or both).

Theorem 4.4. Assume that the subgroup-based protocol is applied. Then every optimal matching is both a RHE and a RUE.

Proof. In the appendix. □

4.4 Synthesis and analysis

We have proposed three cost-sharing protocols to determine how to split the costs among users sharing a ride. All these protocols depend only on the ride itself, without requiring any more information, and two of those are budget-balanced (the remaining one can be made overcharging). Each equilibrium notion from section 3, has at least one corresponding protocol that enables the system to simultaneously reach optimum and equilibrium. Table 1 synthesises which protocol should be used depending on which equilibrium notion governs the system:

Equilibrium	RHE	RUE	RSIE
Protocol to be used	Subgroup-based	Subgroup-based or Residual-based	Externality-based

Table 1: Synthesis of which protocol reaches PoS=1 depending on the equilibrium notion.

Regarding the protocols, the rationale behind each of them can also be synthesised: the externality-based accounts for the extra costs induced to the other members of the group, the residual-based rests mostly on the individual costs (e.g., proportional to the time required to go from the origin to the destination), whereas the subgroup-based recognises which users could be efficiently matched together in smaller groups to determine how to split the costs.

5 Numerical simulations

We now simulate the resulting feasible groups emerging from a real-life case, in order to analyse which matching could emerge for all the equilibrium notions discussed above, and the role played by the cost-sharing protocols we propose.

5.1 The scenario

We consider a batch of 400 travellers departing within a 10-minute window (2400 passengers/h) in the afternoon peak-hour in Amsterdam. Trips, defined through the exact origin, destination and departure time, are sampled from the nation-wide demand dataset (Arentze & H. J. Timmermans 2004).

We assume that the cost $c(G)$ of each group $G \in \mathcal{G}$ comprises users' costs (the monetary equivalent of their time) and operators' monetary costs. For passenger i in group G , we denote by $t(G, i)$ and $w(G, i)$ her in-vehicle travel and waiting times, respectively. We convert time into money using average value of time (9 €/hour) as β_t , and we multiply it by 1.5 to get the penalty for waiting β_w . We assume here that travel time is weighted equally regardless of the number of co-travellers. Therefore, the time-related cost of user i travelling in group G is:

$$C_T(G, i) = \beta_t t(G, i) + \beta_w w(G, i) \quad (36)$$

We calculate the vehicle costs $C_O(G)$ proportional to the trip distance $l(G)$ plus a fixed ride cost:

$$C_O(G) = \beta_l l(G) + \beta_V, \quad (37)$$

Hence the total cost of group G is computed by adding operators' and users' costs:

$$c(G) = \sum_{i \in G} C_T(G, i) + C_O(G) \quad (38)$$

with $\beta_l = 1 \text{ €}/\text{Kilometer}$ and $\beta_V = 1 \text{ €}$.

To compute the set of feasible groups \mathcal{G} we apply the hierarchical exact algorithm of ExMAS (Kucharski & Cats 2020) which computes all feasible groups of travellers (of any size). A group is declared feasible when for all the travellers the additional detour and delay can be compensated thanks to sharing. ExMAS in this configuration generates 2191 feasible groups of various sizes, constituting \mathcal{G} for further calculations.

We first calculate the pricings in the three proposed protocols for each traveller-ride combination. Then we prune the groups according to the different notions of equilibria, and finally we perform the matching to assign travellers to groups. Matching is done first with the objective to minimise total travel costs (to compute the price of stability of the system) and then to maximise it (to compute potential anarchy of the system).

5.2 Results

First we illustrate how the different equilibrium notions restrict the number of feasible solutions, by means of the so-called 'shareability graph' (Figure 2). In such a graph, each traveller is a node, and nodes i and j are connected if the corresponding users could be feasibly matched together (that is, if the group $\{i, j\} \in \mathcal{G}$). In Figure 2 we show how NE and RHE pruned the initial \mathcal{G} (recall that the other notions of equilibrium - RUE and RSIE - prune solutions based on exclusive pairs of groups rather than unfeasible groups). Out of 2191 initially feasible groups, and if using the subgroup-based protocol, 1708 remain for NE and 1366 for RHE. However, such pruning does not significantly alter the graph structure, that always consists of one highly connected giant component and few disconnected nodes.

Table 2 summarises the results of the simulation of the effects of the different equilibrium notions, where we report how the 2191 initially feasible groups are pruned and how it affects the size of the remaining groups. NE prunes mainly groups of size 2, with groups shared by more than two travellers remaining almost intact. Consequently, filtering 2191 groups to 1708 for NE decreases the mean degree mildly. When we allow subgroups to coordinate to leave together, i.e. when we study RHE, rides of greater size are further pruned, resulting in a significantly lower average degree.

We also report in Table 2 the number of mutually exclusive constraints imposed by the RUE and RSIE protocols. This number is much higher for RSIE, i.e., when users can move individually to other group willing to receive her, the resulting equilibria seem to be much more restrictive than when admitting merges between groups. This fits the fact that a RUE always exists, whereas a RSIE does not necessarily exist.

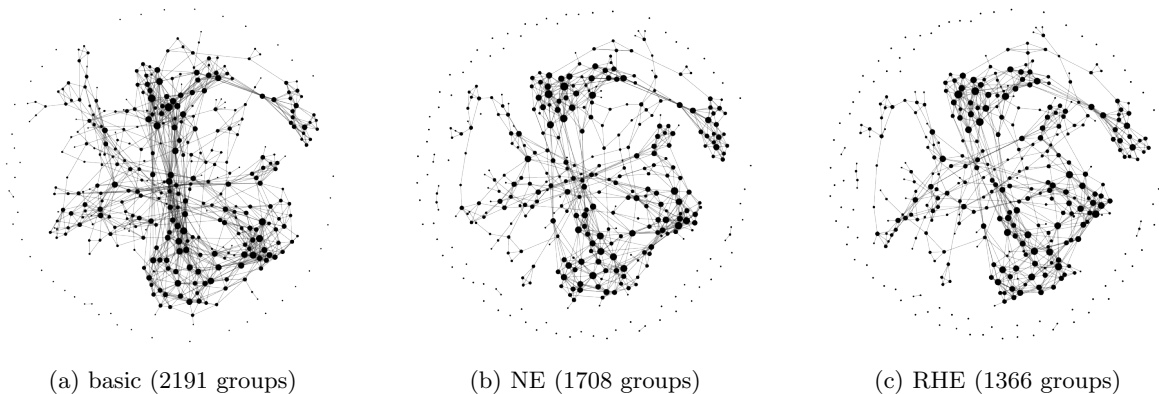


Figure 2: Shareability graphs for selected pruning algorithms. Nodes denote travellers, which are linked if they can share a ride. Nodes are sized according to their degree (i.e. number of connecting edges). The number of nodes remains fixed across the pruning, and the number of links decreases as groups are being excluded in the pruning procedures. Individual costs are defined according to the subgroup-based protocol.

We now analyse the different cost-sharing protocols. Intuitively, one expects that shorter trips result in lower costs, as well as groups of larger sizes (because the costs are split among many). This

algorithm	number of groups	of size				mutually exclusives	mean size
		1	2	3	4+		
Basic	2191	400	1348	363	79	-	2.05
NE	1708	400	928	305	74	-	2.03
HERMETIC	1366	400	806	142	18	-	1.83
RUE	2191	400	1348	363	79	2762	2.05
RSIE	2191	400	1348	363	79	24404	2.05

Table 2: The number of groups and the distribution of their sizes yielded by each of the pruning algorithms, as well as the number of mutually exclusive constrains (between group pairs) in the respective equilibrium notions. Individual costs are defined according to the subgroup-based protocol.

is studied in Figure 3, where we show users’ costs for the 400 travellers in 2191 groups (4085 traveller-group pairs), and their relationship with distance and size, for the different cost-sharing protocols; we also compare these results with users’ direct costs $C_T(G, i)$, i.e., when accounting only for their travelling times.

The top row of Figure 3 shows the relationship between costs and distance. The colour of the dots represents the size of the respective group. We see that direct costs (left panel) are always lower for single rides, because there is no detour. However, when a cost-sharing protocol is introduced, that is, when the operators’ costs are split among the users, larger groups become more attractive and mostly dominate private rides. This conclusion is reinforced when examining the bottom row, that shows users’ costs depending on the groups’ size, where it is apparent that the increasing trend observed in the left panel is neutralised or reversed in the other panels. There are some noteworthy aspects for each protocol:

- **Subgroup-based:** This protocol does not exhibit a high correlation between distance and users’ costs, as many dots are placed far from the diagonal. In addition, it is the protocol in which large groups are favoured the most, as most red dots present the lowest cost for a given distance. This is in line with intuition, as this protocol is meant to ensure that efficient groups are hermetic, so that (a number of) their members are assigned with the lowest costs when the group is complete.
- **Residual-based:** This protocol exhibits the highest correlation between distance and users’ costs, which is expected since costs are calculated proportionally to distances.
- **Externality-based:** This protocol exhibits the lowest correlation between distance and users’ costs. This means that the protocol is effectively capturing that the relevant aspect for this protocol is not the total distance, but the total detour imposed on others. The few dots at the bottom with zero costs represent cases in which a group H would not be feasible according to the ExMAS algorithm which determines the set \mathcal{G} , but there is some feasible group G containing H . In those cases, H is added assuming the cost of the cheapest feasible group containing it.

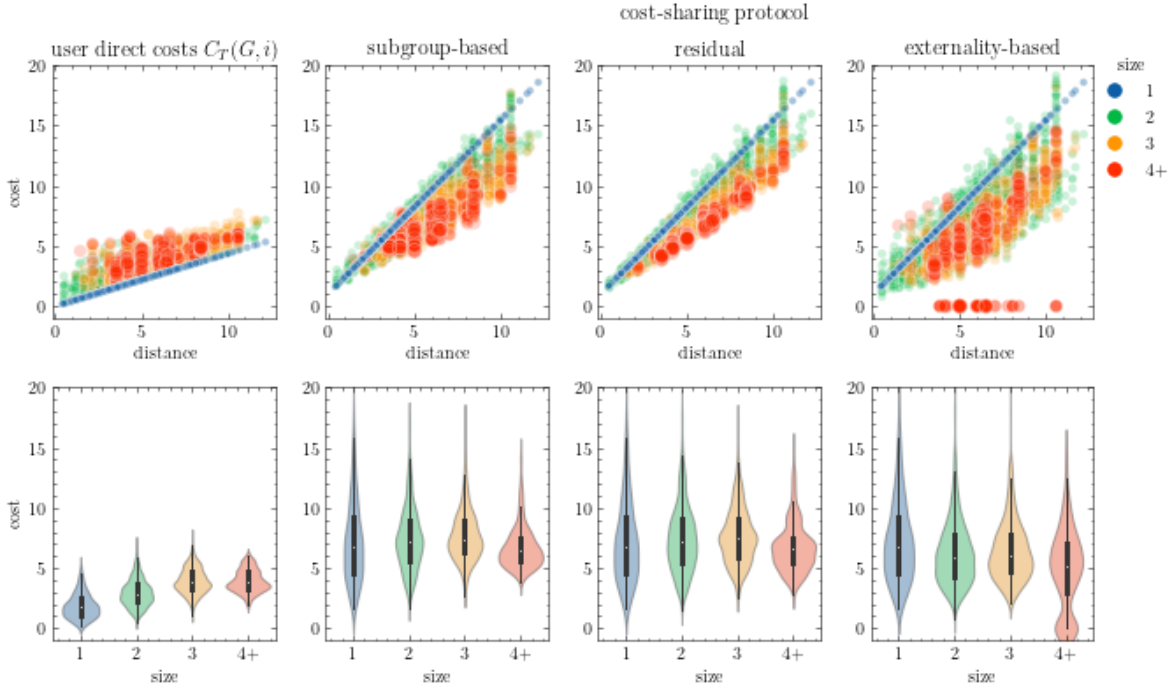


Figure 3: User costs as a function of distance (top row). First we plot direct user costs, followed by the costs under the three introduced cost-sharing protocols. Each dot denotes a traveller-group pair and colours denote the size of the corresponding group. The blue diagonal line denotes non-pooled rides. The bottom row shows the cost distributions varying with the groups' size for the different protocols.

In Table 3 we report the most relevant indices for each protocol. We study the case in which the operator can propose a solution ('Best case', the one that we focus on throughout the paper), but we also analyse the 'Worst case' where we assume the users are not coordinated, so they may reach the solution of maximal, rather than minimal, system-wide costs.

We report passenger-hours, vehicle-hours, number of groups and total cost (reported as PoS - best and PoA - worst, respectively). The best-case analysis confirms that for each equilibrium there is at least one pricing allowing to reach the PoS= 1 (as synthesised in Table 1 in the previous section), through the optimal assignment that requires 234 vehicles, involving 84.01 Pax-Hours and 41.64 Veh-Hours. Moreover, for all the protocols, the price of stability is low regardless of the equilibrium notion, meaning that the protocols are robust. In fact, all the indices are very similar across the protocols, i.e., they yield similar best solutions.

The worst-case results show that the price of anarchy is between 1.1 and 1.22 in this example. That is to say, if the system does not propose a solution, losses can be as high as 22%, which highlights the relevance of having centralised solutions. The residual-based protocol is the most effective one in the case with no central coordination.

Note that the highest PoA is always reached when studying RHE, because everybody travelling alone is a RHE, which is exactly the worst solution (400 groups). For RUE and RSIE the contrary happens, i.e., passenger-hours are similar (albeit higher) to those obtained in the best-case analysis, while increasing vehicle-hours and the number of groups. That is to say, while the system is providing almost the same quality of service as in the best-case scenario, it does so through a non-efficient utilisation of the vehicles. In the case of RUE, this happens because this notion of equilibrium can admit any large group, regardless of its efficiency; something similar happens with RSIE, but indirectly: when groups are large, few individual movements from one group to another are feasible.

In Table 4 we report which portion of the total costs are covered by the sum of users' individual costs in the best-case solution for the externality-based protocol. Even though more than 80% of the total costs are covered by the sum of users' costs, additional subsidies would be needed regardless of the equilibrium notion. It is worth recalling that such need for subsidies can be avoided by means of Corollary 4.2.1, but this would increase the fares for every user. As the other two protocols are budget-balanced, it is guaranteed that the group costs are exactly covered by the individual costs (in other words, the second column of the equivalent to Table 4 would have a "1" in every row).

Protocol	Equilibrium	Best-Case				Worst-Case			
		Pax-Hours	Veh-Hours	N ^o of groups	PoS	Pax-Hours	Veh-Hours	N ^o of groups	PoA
Residual-based	RHE	83.64	41.84	234	1.001	60.34	60.34	400	1.22
	RUE	84.01	41.64	234	1	89.31	54.46	275	1.18
	RSIE	84.12	41.66	234	1.001	93.28	49.08	235	1.12
Subgroup-based	RHE	84.01	41.64	234	1	60.34	60.34	400	1.22
	RUE	84.01	41.64	234	1	88.82	54.57	277	1.19
	RSIE	83.16	42.25	235	1.003	93.4	49.53	234	1.12
Externality-based	RHE	82.14	42.62	240	1.006	60.34	60.34	400	1.22
	RUE	84.01	41.64	234	1	89.02	54.59	276	1.19
	RSIE	84.01	41.64	234	1	92.88	50.11	238	1.13

Table 3: KPIs for the three cost-sharing protocols proposed in the paper, depending on the equilibrium notion that governs the system. On the left side we show the ‘best-case’ results, i.e. when we minimise total costs subject to the respective equilibrium, representing the best solution that could be proposed by the system’s operator. On the right side we show the ‘worst-case’ solution, obtained by maximising costs subject to the respective equilibrium, representing the worst possible outcome if users coordinate by themselves.

Equilibrium	Portion
RHE	0.81
RUE	0.82
RSIE	0.82

Table 4: Portion of the total costs covered by the sum of users’ costs when the externality-based protocol is utilised, considering the best-case solution.

Finally, in Figure 4 we study users’ satisfaction in comparison to the cheapest alternative they have (i.e., if they had the opportunity to select their most preferred group, regardless of their co-travellers’ opinion). To do this, we plot the cumulative distribution of the relative differences between the cheapest alternative for each user, and the individual cost of the group that actually contains her in the best possible matching for each equilibrium notion. Null difference implies a perfect match (best personal option) which increases as users are matched to more expensive pooled rides. These curves complement system-wide indicators, as they allow to investigate distributional effects among users. The faster the curve reaches 1, the more equalised the matching.

We observe that the panels in Figure 4 look similar regardless of the equilibrium notion. For all of them, the externality-based protocol provides the least equal outcome, with some users facing costs that can be higher than 2 times their best option, whereas using the subgroup-based or residual-based protocols this number is reduced to little more than 1.5. The residual-based protocol achieves the highest equity. Recalling that RSIE charges less from the users than the other protocols (as it does not collect enough to cover the whole operators’ costs), an interesting trade-off emerges between total users’ costs and its (unequal) distribution.

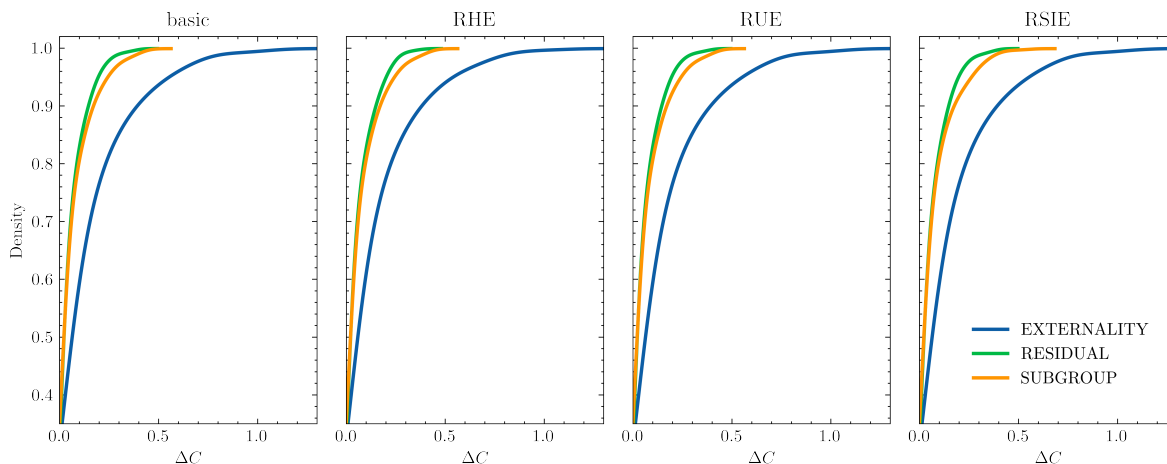


Figure 4: Cumulative distribution of relative differences between each user’s lowest possible cost, and the cost imposed on her by the best possible matching for each equilibrium notion. Different colors represent the various cost-sharing protocols.

6 Synthesis, conclusions and future research

In this paper, we address the issue of how to split common costs when users share a ride in a mobility system that decides how to group the users. After recognising that the way costs are split can affect which groups are going to be formed and hence the quality of the solution (matching), we have modelled the said situation as a game, in which each user can choose with whom to travel, as long as all co-travellers agree.

In order to study the possible equilibria in such a game, we have proved that it suffices to consider pure strategies only. Moreover, we have discussed that the traditional notions of Nash and Strong Equilibria are not the most appropriate ones as they may either prevent sharing on one hand (NE) or require unfeasible coordination on the other (SE). We therefore proposed three intermediate definitions of equilibrium depending on which are the possible ways in which users can coordinate. Moreover, we have also proved that determining whether a Strong Equilibrium exists is an NP-Complete problem, through a reduction from the so-called 3-Exact-Cover.

For each of these equilibrium notions we have proposed a corresponding cost-sharing mechanism that reaches a price of stability equal to one, i.e., that makes any optimal solution also an equilibrium. By this means, we allow a central operator to group the users, simultaneously reaching a system-wide optimum and a users' equilibrium. When deciding what is the cost for an individual within a given group, the protocols determine based on (i) the cost of the respective private trip ('residual-based'), (ii) the extra-costs imposed to the co-travellers ('externality-based'), or (iii) the costs of the subgroups that contain the said user ('subgroup-based').

We tested our ideas by computing the feasible groups for a set of 400 travellers in Amsterdam forming 2191 feasible groups. Numerical results show that our protocols effectively make efficient larger groups to be preferable by the users, and that they always reach a price of stability close to 1. We also show that if the system cannot propose a solution and users coordinate by themselves then the worst case induces around 20% extra-costs for two of the equilibrium notions.

Our discussion and methods demonstrate that selecting an appropriate cost-sharing mechanism, understandable by the users, can play a key role in ensuring that an equilibrium exists and in aligning users' interests with a system-wide optimum. Furthermore, the ability of proposing a central coordination is also crucial to make this type of mobility systems attractive.

As this is an emerging topic, there are numerous directions for further research. In this paper, we have assumed a central operator that aims for a global optimum; if a for-profit company was considered instead, our model would have to be modified to consider the company's interests as well (for instance, as an additional player). We assumed the demand is exogenous and fixed, whereas in fact such cost-sharing protocols can make the system more attractive and thus induce the demand, triggering a positive feedback loop as the critical mass needed for pooling is reached. Moreover, we have assumed that the system's optimum depends only on total costs, regardless of equity aspects, although we show that results can actually be far from equal for all the users involved, which suggests yet another direction for further research. On a theoretical note, studying protocols that ensure bounds on the price of anarchy, as well as determining the complexity of deciding whether a RSIE exists if all the groups have size no larger than 2, are relevant questions that emerge from this paper. Finally, equilibrium analysis in a dynamic environment, taking into account users that will emerge in the future for which there is partial or no information available, is also an intriguing research avenue, that has been analysed only from the point of view optimal matching, both theoretically (Feng et al. 2020; Torrico & Toriello 2017) and in applied models (Fielbaum et al. 2021b; Wallar et al. 2018; Wen et al. 2017; van Engelen et al. 2018; Alonso-Mora et al. 2017a).

Acknowledgments

The authors want to thank José Correa, from Universidad de Chile, and Orlando Rivera, from Universidad Adolfo Ibáñez, for the fruitful discussions that helped to elaborate the contents of this paper. The work of the second and third authors was supported by the CriticalMaaS project (grant number 804469), which is financed by the European Research Council and the Amsterdam Institute for Advanced Metropolitan Solutions.

References

1. Agatz, N., Erera, A., Savelsbergh, M. & Wang, X. Optimization for dynamic ride-sharing: A review. *European Journal of Operational Research* **223**, 295–303 (2012).
2. Albers, S. On the value of coordination in network design. *SIAM Journal on Computing* **38**, 2273–2302 (2009).
3. Alonso-González, M. J., Cats, O., van Oort, N., Hoogendoorn-Lanser, S. & Hoogendoorn, S. What are the determinants of the willingness to share rides in pooled on-demand services? *Transportation*, 1–33 (2020).
4. Alonso-González, M. J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S. & Hoogendoorn, S. Value of time and reliability for urban pooled on-demand services. *Transportation Research Part C: Emerging Technologies* **115**, 102621 (2020).
5. Alonso-Mora, J., Wallar, A. & Rus, D. *Predictive Routing for Autonomous Mobility-on-Demand Systems with Ride-Sharing* in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Sept. 2017), 3583–3590.
6. Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E. & Rus, D. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences* **114**, 462–467 (2017).
7. Anshelevich, E., Dasgupta, A., Kleinberg, J., Tardos, É., Wexler, T. & Roughgarden, T. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing* **38**, 1602–1623 (2008).
8. Arentze, T. A. & Timmermans, H. J. A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological* **38**, 613–633 (2004).
9. Arin, J., Inarra, E. & Luquin, P. A noncooperative view on two airport cost sharing rules. *Review of Economic Design* **13**, 361–376 (2009).
10. Asghari, M., Deng, D., Shahabi, C., Demiryurek, U. & Li, Y. *Price-aware real-time ride-sharing at scale: an auction-based approach* in *Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems* (2016), 1–10.
11. Bian, Z., Liu, X. & Bai, Y. Mechanism design for on-demand first-mile ridesharing. *Transportation Research Part B: Methodological* **138**, 77–117 (2020).
12. Bistaffa, F., Farinelli, A., Chalkiadakis, G. & Ramchurn, S. D. A cooperative game-theoretic approach to the social ridesharing problem. *Artificial Intelligence* **246**, 86–117 (2017).
13. Buchbinder, N., Lewin-Eytan, L., Naor, J. S. & Orda, A. *Non-cooperative cost sharing games via subsidies* in *International Symposium on Algorithmic Game Theory* (2008), 337–349.
14. Caragiannis, I., Gkatzelis, V. & Vinci, C. *Coordination mechanisms, cost-sharing, and approximation algorithms for scheduling* in *International Conference on Web and Internet Economics* (2017), 74–87.
15. Castillo, J. C., Knoepfle, D. & Weyl, G. *Surge pricing solves the wild goose chase* in *Proceedings of the 2017 ACM Conference on Economics and Computation* (2017), 241–242.
16. Chan, N. D. & Shaheen, S. A. Ridesharing in North America: Past, present, and future. *Transport reviews* **32**, 93–112 (2012).
17. Chau, C. & Elbassioni, K. Quantifying inefficiency of fair cost-sharing mechanisms for sharing economy. *IEEE Transactions on Control of Network Systems* **5**, 1809–1818 (2017).
18. Chau, S. C.-K., Shen, S. & Zhou, Y. Decentralized Ride-Sharing and Vehicle-Pooling Based on Fair Cost-Sharing Mechanisms. *IEEE Transactions on Intelligent Transportation Systems* (2020).
19. Chen, H., Roughgarden, T. & Valiant, G. Designing network protocols for good equilibria. *SIAM Journal on Computing* **39**, 1799–1832 (2010).
20. Chen, L., Zhong, Q., Xiao, X., Gao, Y., Jin, P. & Jensen, C. S. *Price-and-time-aware dynamic ridesharing* in *2018 IEEE 34th International Conference on Data Engineering (ICDE)* (2018), 1061–1072.
21. Cheng, S.-F., Nguyen, D. T. & Lau, H. C. *Mechanisms for arranging ride sharing and fare splitting for last-mile travel demands* in *25th International Joint Conference on Artificial Intelligence* (2014), 475–481.

22. Christodoulou, G., Gkatzelis, V., Latifian, M. & Sgouritsa, A. *Resource-Aware Protocols for Network Cost-Sharing Games* in *Proceedings of the 21st ACM Conference on Economics and Computation* (2020), 81–107.
23. Christodoulou, G., Gkatzelis, V. & Sgouritsa, A. *Cost-sharing methods for scheduling games under uncertainty* in *Proceedings of the 2017 ACM Conference on Economics and Computation* (2017), 441–458.
24. Chvatal, V. A greedy heuristic for the set-covering problem. *Mathematics of operations research* **4**, 233–235 (1979).
25. Cremer, H., Gahvari, F. & Ladoux, N. Externalities and optimal taxation. *Journal of Public Economics* **70**, 343–364 (1998).
26. Diao, M., Kong, H. & Zhao, J. Impacts of transportation network companies on urban mobility. *Nature Sustainability*, 1–7 (2021).
27. Dosa, G. The Intermediate Price of Anarchy (IPoA) in bin packing games. *Discrete Applied Mathematics* **242**, 16–25 (2018).
28. Enzi, M., Parragh, S. N., Pisinger, D. & Prandtstetter, M. Modeling and solving the multimodal car-and ride-sharing problem. *European Journal of Operational Research* **293**, 290–303 (2021).
29. Escoffier, B., Gourves, L. & Monnot, J. *On the impact of local taxes in a set cover game* in *International Colloquium on Structural Information and Communication Complexity* (2010), 2–13.
30. Feng, Y., Niazadeh, R. & Saberi, A. Two-stage Matching and Pricing with Applications to Ride Hailing. *Chicago Booth Research* **20-28** (2020).
31. Fielbaum, A. & Alonso-Mora, J. Unreliability in ridesharing systems: Measuring changes in users’ times due to new requests. *Transportation Research Part C: Emerging Technologies* **121**, 102831 (2020).
32. Fielbaum, A., Bai, X. & Alonso-Mora, J. On-demand ridesharing with optimized pick-up and drop-off walking locations. *Transportation Research Part C: Emerging Technologies* **126**, 103061 (2021).
33. Fielbaum, A., Kronmueller, M. & Alonso-Mora, J. Anticipatory routing methods for an on-demand ridepooling mobility system. *Transportation*, 1–42 (2021).
34. Frank, M. The braess paradox. *Mathematical Programming* **20**, 283–302 (1981).
35. Furuhata, M., Daniel, K., Koenig, S., Ordonez, F., Dessouky, M., Brunet, M.-E., Cohen, L. & Wang, X. Online cost-sharing mechanism design for demand-responsive transport systems. *IEEE Transactions on Intelligent Transportation Systems* **16**, 692–707 (2015).
36. Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M.-E., Wang, X. & Koenig, S. Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B: Methodological* **57**, 28–46 (2013).
37. Gairing, M., Kollias, K. & Kotsialou, G. Existence and efficiency of equilibria for cost-sharing in generalized weighted congestion games. *ACM Transactions on Economics and Computation (TEAC)* **8**, 1–28 (2020).
38. Garey, M. R. & Johnson, D. S. *Computers and intractability* (Freeman San Francisco, 1979).
39. Gkatzelis, V., Kollias, K. & Roughgarden, T. Optimal cost-sharing in general resource selection games. *Operations Research* **64**, 1230–1238 (2016).
40. Gkatzelis, V., Kollias, K. & Roughgarden, T. *Optimal cost-sharing in weighted congestion games* in *International Conference on Web and Internet Economics* (2014), 72–88.
41. Gottlob, G., Greco, G. & Scarcello, F. Pure Nash equilibria: Hard and easy games. *Journal of Artificial Intelligence Research* **24**, 357–406 (2005).
42. Gusfield, D. The structure of the stable roommate problem: efficient representation and enumeration of all stable assignments. *SIAM Journal on Computing* **17**, 742–769 (1988).
43. Harks, T. & Miller, K. The worst-case efficiency of cost sharing methods in resource allocation games. *Operations research* **59**, 1491–1503 (2011).
44. Harks, T. & Timmermans, V. Equilibrium computation in resource allocation games. *Mathematical Programming*, 1–34 (2021).

45. Harks, T. & von Falkenhausen, P. Optimal cost sharing for capacitated facility location games. *European Journal of Operational Research* **239**, 187–198 (2014).
46. Henao, A. & Marshall, W. E. The impact of ride-hailing on vehicle miles traveled. *Transportation* **46**, 2173–2194 (2019).
47. Hernández, R., Cárdenas, C. & Muñoz, D. Game theory applied to transportation systems in Smart Cities: analysis of evolutionary stable strategies in a generic car pooling system. *International Journal on Interactive Design and Manufacturing (IJIDeM)* **12**, 179–185 (2018).
48. Ho, C. Q., Hensher, D. A., Mulley, C. & Wong, Y. Z. Potential uptake and willingness-to-pay for Mobility as a Service (MaaS): A stated choice study. *Transportation Research Part A: Policy and Practice* **117**, 302–318 (2018).
49. Hu, S., Dessouky, M. M., Uhan, N. A. & Vayanos, P. Cost-sharing mechanism design for ride-sharing. *Transportation Research Part B: Methodological* **150**, 410–434 (2021).
50. Jacob, J. & Roet-Green, R. Ride solo or pool: Designing price-service menus for a ride-sharing platform. *European Journal of Operational Research* (2021).
51. Jain, K. & Mahdian, M. Cost sharing. *Algorithmic game theory* **15**, 385–410 (2007).
52. Jara-Díaz, S. *Transport economic theory* (Elsevier Oxford, UK, 2007).
53. Ke, J., Yang, H., Li, X., Wang, H. & Ye, J. Pricing and equilibrium in on-demand ride-pooling markets. *Transportation Research Part B: Methodological* **139**, 411–431 (2020).
54. Kleiner, A., Nebel, B. & Ziparo, V. *A mechanism for dynamic ride sharing based on parallel auctions* in *Twenty-Second International Joint Conference on Artificial Intelligence* (2011).
55. Kucharski, R. & Cats, O. Exact matching of attractive shared rides (ExMAS) for system-wide strategic evaluations. *Transportation Research Part B: Methodological* **139**, 285–310 (2020).
56. Kucharski, R., Fielbaum, A., Alonso-Mora, J. & Cats, O. If you are late, everyone is late: late passenger arrival and ride-pooling systems’ performance. *Transportmetrica A: Transport Science*, 1–24 (2020).
57. Levinger, C., Hazon, N. & Azaria, A. *Computing the Shapley Value for Ride-Sharing and Routing Games* in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (2020), 1895–1897.
58. Lu, W. & Quadrioglio, L. Fair cost allocation for ridesharing services—modeling, mathematical programming and an algorithm to find the nucleolus. *Transportation Research Part B: Methodological* **121**, 41–55 (2019).
59. Ma, K., Wang, C., Yang, J., Hua, C. & Guan, X. Pricing mechanism with noncooperative game and revenue sharing contract in electricity market. *IEEE transactions on cybernetics* **49**, 97–106 (2017).
60. Mogridge, M. J. The self-defeating nature of urban road capacity policy: A review of theories, disputes and available evidence. *Transport Policy* **4**, 5–23 (1997).
61. Moulin, H. The price of anarchy of serial, average and incremental cost sharing. *Economic Theory* **36**, 379–405 (2008).
62. Moulin, H. Axiomatic cost and surplus sharing. *Handbook of social choice and welfare* **1**, 289–357 (2002).
63. Mourad, A., Puchinger, J. & Chu, C. A survey of models and algorithms for optimizing shared mobility. *Transportation Research Part B: Methodological* **123**, 323–346 (2019).
64. Nisan, N. & Ronen, A. Computationally feasible VCG mechanisms. *Journal of Artificial Intelligence Research* **29**, 19–47 (2007).
65. Özkan, E. Joint pricing and matching in ride-sharing systems. *European Journal of Operational Research* **287**, 1149–1160 (2020).
66. Peng, Z., Shan, W., Jia, P., Yu, B., Jiang, Y. & Yao, B. Stable ride-sharing matching for the commuters with payment design. *Transportation* **47**, 1–21 (2020).
67. Rasulkhani, S. & Chow, J. Y. Route-cost-assignment with joint user and operator behavior as a many-to-one stable matching assignment game. *Transportation Research Part B: Methodological* **124**, 60–81 (2019).

68. Raz, R. & Safra, S. *A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP* in *STOC* **97** (1997), 475–484.
69. Roughgarden, T. & Schrijvers, O. Network cost-sharing without anonymity. *ACM Transactions on Economics and Computation (TEAC)* **4**, 1–24 (2016).
70. Roughgarden, T. & Sundararajan, M. Quantifying inefficiency in cost-sharing mechanisms. *Journal of the ACM (JACM)* **56**, 1–33 (2009).
71. Schroder, M., Storch, D.-M., Marszal, P. & Timme, M. Anomalous supply shortages from dynamic pricing in on-demand mobility. *Nature Communications* **11** (2020).
72. Shen, W., Lopes, C. V. & Crandall, J. W. An online mechanism for ridesharing in autonomous mobility-on-demand systems. *arXiv preprint arXiv:1603.02208* (2016).
73. Tirachini, A. & Gomez-Lobo, A. Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. *International journal of sustainable transportation* **14**, 187–204 (2020).
74. Torrico, A. & Toriello, A. Dynamic Relaxations for Online Bipartite Matching. *arXiv preprint arXiv:1709.01557* (2017).
75. Van Engelen, M., Cats, O., Post, H. & Aardal, K. Enhancing flexible transport services with demand-anticipatory insertion heuristics. *Transportation Research Part E: Logistics and Transportation Review* **110**, 110–121 (2018).
76. Von Falkenhausen, P. & Harks, T. Optimal cost sharing for resource selection games. *Mathematics of Operations Research* **38**, 184–208 (2013).
77. Von Falkenhausen, P. & Harks, T. *Optimal cost sharing protocols for scheduling games* in *Proceedings of the 12th ACM conference on Electronic Commerce* (2011), 285–294.
78. Wallar, A., Van Der Zee, M., Alonso-Mora, J. & Rus, D. *Vehicle rebalancing for mobility-on-demand systems with ride-sharing* in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2018), 4539–4546.
79. Wang, X., Agatz, N. & Erera, A. Stable matching for dynamic ride-sharing systems. *Transportation Science* **52**, 850–867 (2018).
80. Wen, J., Zhao, J. & Jaillet, P. *Rebalancing shared mobility-on-demand systems: A reinforcement learning approach* in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (2017), 220–225.
81. Wu, X. & MacKenzie, D. Assessing the VMT effect of ridesourcing services in the US. *Transportation Research Part D: Transport and Environment* **94**, 102816 (2021).
82. Yan, P., Lee, C.-Y., Chu, C., Chen, C. & Luo, Z. Matching and Pricing in Ride-Sharing: Optimality, Stability, and Financial Sustainability. *Omega*, 102351 (2020).
83. Zhang, H. & Zhao, J. Mobility sharing as a preference matching problem. *IEEE Transactions on Intelligent Transportation Systems* **20**, 2584–2592 (2018).
84. Zhang, J., Wen, D. & Zeng, S. A discounted trade reduction mechanism for dynamic ridesharing pricing. *IEEE Transactions on Intelligent Transportation Systems* **17**, 1586–1595 (2015).

Appendix

Continuation of the proof of Theorem 3.4 a)

We now need to prove that there is an exact cover in the original instance of 3-exact-cover iff there is a strong equilibrium in this instance of *CTG*.

If there is an exact cover b_1, \dots, b_p in the original instance, it is straightforward how to build a strong equilibrium in the instance of *CTG*. Indeed, consider the profile of strategies in which α travels alone, and the other groups are b_1, \dots, b_p . Therefore, every element of S is in its most preferred situation, and α has no other choice than being alone.

We will now show that these are the only possible strong equilibria in such an instance of *CTG*, more precisely we will prove the following Lemma:

Lemma 6.1. *Let h_1, \dots, h_q be a strong equilibrium on the proposed instance of *CTG*. Then there exists ℓ such that $h_\ell = \{\alpha\}$ and $\forall j \neq \ell, h_j$ corresponds to some b_i from \mathcal{B} .*

Lemma 6.1 implies that when there is a strong equilibrium in the instance of CTG , we can construct the exact cover just by taking all such subsets h_j . The proof of Lemma 6.1 rests on another technical Lemma, that limits which situations can occur in a strong equilibrium:

Lemma 6.2. *Let h_1, \dots, h_q be a strong equilibrium on the proposed instance of CTG . Then none of the following situations can happen:*

- i) *The existence of one h_i of size 3 that does not come from \mathcal{B} and the existence of one h_j of size one.*
- ii) *The existence of two h_i, h_j both of size 2.*
- iii) *The existence of two h_i, h_j both of size 1.*

Before proving Lemma 6.2, let us explain why it suffices to prove Lemma 6.1. In fact, the number of players in this game is $3t + 1$, for some $t \in \mathbb{N}$. If we remove all those players that belong to a subset h_i coming from \mathcal{B} , the remaining number of players is $3s + 1$ for some $s \in \mathbb{N}$. We just need to show that $s = 0$, i.e., that all players but α are covered by groups coming from \mathcal{B} . What Lemma 6.2 ensures is that if $s > 0$, then it would be impossible to cover the remaining $3s + 1$ players with the subsets h_i not coming from \mathcal{B} , which is a contradiction. Hence, Lemma 6.1 is proven by Lemma 6.2.

Putting everything together, we just need to prove that Lemma 6.2 is true. We will show that each of the forbidden situations is indeed impossible in a strong equilibrium, following the same order:

- i) Assume that there is a group $h_i = \{x, y, z\}$ that does not come from \mathcal{B} , and a group $h_j = \{w\}$. Without loss of generality, w falls between x and y within the circle that defines the lexicographic order of the preferences. Hence, both x and w would be strictly better if they change to form the group $\{x, w\}$, contradicting the fact that this was a strong equilibrium.
- ii) Assume two groups $h_i = \{x, y\}, h_j = \{w, z\}$. When we restrict the circle to these four players, either the preferred co-player for x is not y , or the preferred co-player for y is not x (or both). Without loss of generality, we assume the former case. It is clear that the users of a group of size 2 will always improve their situation if someone else joins, then x, z and w would coordinate to form the group $\{x, z, w\}$, contradicting the fact that this was a strong equilibrium.
- iii) Two isolated players will always prefer to merge.

I.e., the forbidden situations from Lemma 6.2 cannot exist in a strong equilibrium, which completes the proof the Theorem.

□

Continuation of the proof of Theorem 3.4 b)

The stable-roommate problem (SRP) is defined by a set of 2μ players for some $\mu \in \mathbb{N}$, such that for each player there is a list sorting the rest of the players in some strict order of preferences. The problem consists on determining whether all the players can be grouped in pairs, in a way where it never happens that i and j are not together but they both would be better if they were. We now prove that SRP is equivalent to determining the existence of a SE in the restricted version of CTG where all groups have size 1 or 2, that we denote $CTG - 2$.

Proof. Let \mathcal{G} be the set of groups in an instance of $CTG - 2$. Consider a user $i \in P$. Note that her preferences can be described as a list $L_i = (y_1, \dots, y_{k_i}, i)$, meaning that her preferred group is (i, y_1) , followed by (i, y_2) , and so on until (i, y_{k_i}) and then to travel alone. The groups that come after travelling alone are not relevant as they shall never be part of any equilibrium. Utilising this notation, we build an instance of SRP as follows:

- The set of players is $P \times P'$, where P' contains one copy of each player in P .
- If the list of preferences for $x \in P$ is $L_x = (y_1, \dots, y_{k_x}, x)$ in $CTG - 2$, the preferences in SRP are $(y_1, \dots, y_{k_x}, x')$ (i.e., the same one but switching x by its copy in P'); the order of the players after x' is irrelevant for this proof.
- For $x' \in P'$, its preferred match would be $x \in P$. The list continues with the elements in P' according to a circle (i.e., it prefers y' over z' iff $y' - x' \bmod n < z' - x' \bmod \mu$, as described in the proof of the part a) of this Theorem), and the rest of the preferences is irrelevant.

We need to prove that there is a strong equilibrium in the original instance of $CTG - 2$ iff there is a stable matching in this instance of SRP . To do that, it is useful to note first that an even-sized circle $\{x_1, \dots, x_{2\mu}\}$ always presents a stable matching, by joining x_i with $x_{i+\mu \pmod{2\mu}}$.

Take first a strong equilibrium in $CTG - 2$, formed by the pairs $(a_1, b_1), \dots, (a_k, b_k)$ and the users c_1, \dots, c_ℓ travelling alone. We build a stable matching in SRP as follows:

- a_i and b_i are matched in P for all $i = 1, \dots, k$.
- c_i in P is matched with c'_i in P' for all $i = 1, \dots, \ell$.
- The only players that remain to be matched are the copies in P' of a_i and b_i , for all $i = 1, \dots, k$. Thus, it is an even number of players in P' , that form a circle, so we can create a stable matching among them as discussed above.

It is straightforward to see that this is indeed a stable matching, thanks to the fact that it was built from a strong equilibrium in $CTG - 2$.

Take now a stable matching in SRP . We first note that in this stable matching there is no pair (x, y') , with $x \in P, y' \in P'$ and $x \neq y'$: if there was such a pair, then the matching would not be stable because x and x' would prefer be matched together. Therefore, each $x \in P$ is either matched with a different $z \in P$ or with x' . We build the strong equilibrium taking exactly these assignments: each pair (x, z) formed by two elements in P will form a group in $CTG - 2$, and those y that are paired with their copy $y' \in P'$ will travel alone in $CTG - 2$. It is straightforward to see that such an assignment makes a strong equilibrium, thanks to the fact that it was built from a stable matching. \square

Proof of Theorem 3.5

We prove the Theorem by finding a RUE, which is done algorithmically. In short, we begin with everybody travelling alone, and each time we find two mergeable groups, we merge them, which we repeat until we find no more. The resulting profile of strategies is a RUE. We provide the respective pseudo-code in algorithm 4.

Algorithm 4: Construction of a RUE.

```

 $\forall i \in P, G_i = \{i\};$ 
 $v = 0;$  %  $v$  is an auxiliary variable to end the following while cycle.
while  $v=0$  do
   $v = 1;$ 
  for all  $i, j \in P$  such that  $G_i \neq G_j$  do
    if  $G_i$  and  $G_j$  are mergeable then
       $v = 0;$  % The cycle continues
       $G_i, G_j \leftarrow G_i \cup G_j;$ 
    end if
  end for
end while
Output  $(G_i)_{i \in P};$ 

```

By construction, the output from algorithm 4 is a RUE:

- The algorithm only stops if there are no more pairs of mergeable groups.
- The output is also a NE. As the algorithm only induces Pareto-improvements, everybody ends-up better-off (or equal) than in the initial situation, i.e., than travelling alone. This is exactly the definition of being a NE.

Moreover, algorithm 4 does stop because it starts with a profile of strategies, and at each step that is kept but with groups of increasing size.

Example of a RSIE that is not RUE nor RHE

Consider a game with six players. Every combination of players is feasible. The only thing that matters for the players is the size of the group that they belong in, with the following order (from best to worse): 6, 2, 5, 3, 1, 4. In such a setting, splitting the six players into two groups of size 3 constitutes a RSIE that is not RHE nor RUE:

- It is RSIE because any individual movement would lead to a group of size 1 or 4, than are worse than the current situation.
- It is not RUE because the groups could merge and everybody would agree.
- It is not RHE because any pair of players within a group would choose to leave the group to travel together as a pair.

Example of an instance of CTG with no matching that is RHE and RUE

Consider a game with five players A, B, C, D, E . As in the proof of Theorem 3.4, we do not need to impose that if $H \subseteq G$, then $c(H) \leq c(G)$, and it suffices to exhibit the players' orders of preferences. The set \mathcal{G} is defined by:

$$\mathcal{G} = \{A, B, C, D, E, AB, AE, BC, CD, DE, ABC, ABE, ADE, BCD, CDE\}$$

The set \mathcal{G} should have the property that if $G \in \mathcal{G}$ and $H \subseteq G$, then $H \in \mathcal{G}$, which is not the case with this definition of \mathcal{G} . We assume that those groups H are in \mathcal{G} , but all members of H would prefer to travel alone, so H will not be a part of any equilibrium and we can omit them.

In Table 5 we show, for each user, how they sort their preferences. It is assumed that travelling alone comes right after the last element of the respective columns. For instance, the best choice for player A would be to form the group ABC , whereas her worst option before travelling alone is ADE . Note that the preferences are symmetric: for player p_i and using the sum mod 5, her order of preferences is: $p_{i+1}p_{i+2}, p_{i+1}, p_{i+1}p_{i-1}, p_{i-1}, p_{i-2}p_{i-1}$.

A	B	C	D	E
ABC	BCD	CDE	DAE	EAB
AB	BC	CD	DE	EA
ABE	BAC	CBD	DCE	EAD
AE	BA	CB	DC	ED
ADE	BAE	CAB	DBC	ECD
A	B	C	D	E

Table 5: Feasible co-travellers for each of the five users, sorted according to their preferences.

We now show that there is no matching for this instance that is a RUE and a RHE. First, such a matching could not contain any group of size 3 because they are not hermetic. In fact, all those groups are of the type $p_i p_{i+1} p_{i+2}$, and the subgroup $p_{i+1} p_{i+2}$ would always want to leave. Discarding the groups of size 3, there are three remaining options for A : travelling alone, with B or with E . It is useful to note that two groups of the form $p_{i-1} p_i$ and p_{i+1} , respectively, are always mergeable.

- If A travels alone, B can travel alone, but it would merge with A , or B can travel with C . In the last case, D would travel with E , but A and DE are mergeable.
- If A travels with B , C can travel alone or with D . If C travels alone, AB and C are mergeable. If C travels with D , E travels alone, but CD and E are mergeable.
- If A travels with E , B can travel alone or with C . In the first case, AE and B are mergeable. In the second case, BC and D are mergeable.

Therefore, any feasible matching for this instance contains either a pair of mergeable groups or a non-hermetic group.

Proof of Theorem 4.4

The proof consists of two parts. First, we prove that any group with negative excess is hermetic. Second, we show that in any optimal matching every group leads to a negative excess. The fact that every optimal matching is a RUE follows directly from Lemma 4.1.

Let G be a group with negative excess, implying that $c_i(G) \leq z_i(G)$ (Eq. 35), and let $H \subseteq G$. We shall show that users in H do not want to coordinate to leave G . Take $i \in H$ such that

$$\forall j \in H, z_i(G) \leq z_j(G) \tag{39}$$

By definition of the functions z , Eq. (39) implies that $\varphi_i(G)$ was the first one selected among $\{\varphi_j(G) : j \in H\}$. Therefore, the set $\varphi_i(H)$ was completely available (i.e., in W in Algorithm 3) when $\varphi_i(G)$ was selected. This implies that $z_i(H) \geq z_i(G)$, as $z_i(H)$ is computed optimising over a smaller set. If $e(H) \geq 0$, then

$$c_i(H) \geq z_i(H) \geq z_i(G) \geq c_i(G) \quad (40)$$

The positive excess of H explains the first inequality, the second inequality was explained in the previous paragraph and the third inequality is due to the negative excess of G . If $e(H) < 0$ and $c_i = z_i(H)$, Eq. (40) still holds. Finally, if $e(H) < 0$ and $c_i(H) < z_i(H)$:

$$c_i(H) \geq \frac{c(H)}{|H|} \geq z_i(G) \geq c_i(G) \quad (41)$$

Where the first inequality is due to the definition of the costs in the case of negative excess, which diminishes the costs of some subgroups but never below the average cost of the whole group (Eq. 34), the second inequality is explained because $\varphi_i(G)$ is selected when the whole subset H is available, and the third inequality holds because $e(G) \leq 0$ (Eq. 35). Putting everything together, we have shown that $c_i(H) \geq c_i(G)$, i.e., i would not leave G to form H . Therefore, G is indeed hermetic.

The negative excess of G in an optimal matching emerges when recalling that sets φ^j are a partition of G . Indeed:

$$\sum_{i \in G} z_i(G) = \sum_{i \in G} \frac{c(\varphi_i(G))}{|\varphi_i(G)|} = \sum_{j=1}^Q c(\varphi^j) \geq c(G) \quad (42)$$

Note that the third sum is adding each $\varphi_i(G)$ once. The second equality is achieved by noting that each $\frac{c(\varphi_i(G))}{|\varphi_i(G)|}$ is added exactly $|\varphi_i(G)|$ times, and the final inequality is true because sets φ^j are a partition, that has to be sub-optimal because forming the group G is part of the optimal matching.