Getting Your Fare Share

Edgar Chen[1], Omer Lev[2], Ariel D. Procaccia[1], Xinyu Wu[1], and Yair Zick[3]

[1] Carnegie Mellon University, Pennsylvania
{edgarc,xinyuw1}@andrew.cmu.edu,arielpro@cs.cmu.edu
[2] University of Toronto, Canada
omerl@cs.toronto.edu
[3] National University of Singapore
dcsyaz@nus.edu.sg

Abstract. In this work we study several cost allocation methods and their implication in the ridesharing domain. We first identify a list of desiderata: properties (or *axioms*) that any reasonable ridesharing mechanism should satisfy; working from these axioms, we derive several ridesharing mechanisms. Our main technical contribution is the identification of a *unique* strategyproof cost-sharing mechanism, known in the literature as the *constrained equal gains mechanism*, or CEG. Strategyproofness might be desirable in large markets, as it prevents market inefficiency caused by unnecessary strategic behavior on the passengers' side. However, as our simulations show, this comes at some cost: CEG places an unusually high financial burden on players with low costs, and (arguably) over-rewards players with high costs. We show this via a comparison with other cost-sharing mechanisms in simulated environments based on the NYC taxi dataset.

1 Introduction

Cost sharing problems (also known as *bankruptcy problems*) are an age-old problem: a group of people each have a personal cost; however, their joint cost is not the sum of their costs, so a way is needed to divide it between them[4]. This simple abstraction is surprisingly powerful, describing scenarios such as allocating liquidated asset revenue to creditors or dividing cab fares amongst friends. Recent years have seen the rise of a novel application of the bankruptcy problem, arguably exceeding prior ones in both ubiquity and scale. Ridesharing services such as Uber and Lyft have revolutionized the on-demand transportation market, offering a convenient, low-cost alternative to urban commuters. These applications have recently begun to feature *ridesharing services*, where passengers are offered discounted rides if they agree to commute with others. Much like bankruptcy problems, users have an individual cost — how much they would pay had they commuted alone — and the joint cost of the shared ride.

[4] This is inspired by bankruptcies, in which each creditor incurs a cost (the debt), but can recover it only from the estate, which is smaller than the sum of all creditor claims.

The ridesharing problem differs from classic bankruptcy problems in one significant way: ride costs are derived from an underlying graph structure. This raises interesting problems: one is an optimization issue — finding the optimal allocation of riders to shared services is (perhaps unsurprisingly) computationally intractable. However, the underlying graphical model has interesting implications to agents' *strategic behavior*, adding manipulation methods to those commonly seen in bankruptcy problems. In the cost sharing setting, instead of reporting a cost c_i, agent i can misreport her location to be ε far from her true destination, and ride solo the rest of the way. A cost sharing mechanism is manipulable if misreporting is beneficial in some circumstances. By offering users an opportunity to exploit the market for their own gain, manipulable mechanisms risk having commuters strategically gaming the system; indeed, with the increasing proliferation of on-demand ridesharing and self-driving cars becoming a reality, improperly designed fare-sharing mechanisms could result in large-scale market inefficiency.

1.1 Our Contributions

We explore axiomatic fair division methods for ridesharing. We show that certain natural cost sharing axioms are incompatible, while others result in natural mechanisms (such as proportional allocation). More importantly, we study a particular fare allocation method (known in the literature as constrained equal gains, or CEG) which uniquely satisfies a set of natural axioms, as well as strategyproofness. Our proofs require a careful handling of the underlying graph structure; this is despite the fact that our ridesharing mechanisms do not explicitly take the underlying road network as input, as such mechanisms would seem to be quite unwieldy for practical use.

Finally, we demonstrate our result on the New-York City taxi dataset, and in the course of this, we demonstrate that other mechanisms, while not having CEG's axiomatic properties, have some other desirable features. Furthermore, we develop an experimental framework for testing ridesharing mechanisms using Google Maps[5].

1.2 Related Work

Much of the previous work on the ridesharing problem focuses on optimally matching riders to taxis [2], or planning optimal drop-off routes [8, 13, 14]; these problems are computationally hard in general (see [1] for a review). In their review of ridesharing problems, Furuhata et al. [7] mention three pertinent issues: incentivizing truthful behavior from participants, dividing cost fairly, and efficiently handling online rideshare requests. In this paper, we tackle the first two of these issues, effectively showing that egalitarianism (all people pay the same) is deeply linked to strategyproofness, while other notions of fairness are, to a degree, incompatible. Kamar and Horvitz [11] and Kleiner et al. [12] investigate

[5] To maintain blind review, we will add the URL after the review process is complete.

VCG-based mechanisms for real-time cost sharing [10]; these mechanisms focus on real-time ridesharing and do not take an axiomatic approach. Zhao et al. [17] use a somewhat different model of ridesharing, and attempt to use VCG in ridesharing settings, but they find it on satisfactory, resorting to other, less efficient mechanisms. Bistaffa et al. [4] propose a coalitional game approach for finding a fair cost sharing allocation in the ridesharing domain, but their approach is not axiomatic either.

Frisk et al. [6] discuss the cost division problem in a static setting. They axiomatically investigate cost sharing methods like proportional sharing, the Shapley value [16] and Equal Profit Sharing, but find that these can result in unstable allocations. Axiomatic cost sharing is well explored (see [15] for a detailed review), and is often applied to resource or goods allocation [3]. Beyond their applications in ridesharing services, some online services have implemented cost sharing mechanisms for allocating taxi fares. One notable example is the *Spliddit* website (`spliddit.org`) [9], which offers a Shapley value based fare division mechanism. However, as noted by the designers, their mechanism is not even individually rational in some cases.

2 Preliminaries

We are given a set $N = \{1, \dots, n\}$ of players (the passengers), assumed to start from the same location s.[6] Each player $i \in N$ can opt to not use the ridesharing service, and commute alone to her destination at a cost of c_i; we assume that $c_1 \le \cdots \le c_n$. If players choose to share their ride, they incur a total cost of F, where $\sum_{i=1}^{n} c_i \ge F$. Players' costs are induced by an underlying graph structure: the location s is a node in a weighted, directed graph, the costs c_i are the shortest paths to a node v_i which is player i's destination, and the shared cost F is the shortest path passing through the nodes $v_1, \dots, v_n$. Computing F is, rather unsurprisingly, computationally intractable; however, since this work studies the cost sharing problem, rather than the underlying optimization that generates it, we assume that F is given to us as input.

More formally, we are interested in *cost sharing mechanisms*; these are functions whose input is a cost sharing problem $\langle \boldsymbol{c}; F \rangle$, and whose output is a vector $\langle p_1(\boldsymbol{c}; F), \dots, p_n(\boldsymbol{c}; F) \rangle$, where $p_i(\boldsymbol{c}; F)$ is the amount to be paid by player i.

Note that our cost sharing mechanisms do not receive the underlying graph as input. In particular, they do not consider parameters such as the marginal cost of players to the ride. This is a conscious modeling choice in this work: our results can be easily applied in other settings beyond the ridesharing domain; it produces a simple and easily predictable solution to the problem at hand, which can be more easily used (one doubts passengers will relish checking each others'

[6] Since the objective of this work is cost sharing rather than the underlying optimization problem, assuming that all passengers start from the same location does not significantly affect our results (particularly the impossibility results). Practically speaking, though, such a setting arises following well-attended events, such as sports games, when people wish to get home from the event venue.

destination coordinates to determine if a ride is beneficial for them); and it does not necessitate an exponential number of calculations of an intractable problem.

Let us start our investigation by first describing the axioms used in this work.

Efficiency (Eff) Payments sum up to the cost: $\sum_{i=1}^{n} p_i(\boldsymbol{c}; F) = F$.

Symmetry (Sym) If $c_i = c_j$, then $p_i(\boldsymbol{c}; F) = p_j(\boldsymbol{c}; F)$.

Individual Rationality (IR) No agent pays more for ridesharing than they would pay on their own: for all $i \in N$, $p_i(\boldsymbol{c}; F) \leq c_i$.

Non negativity (NN) $p_i(\boldsymbol{c}; F) \geq 0$.

Strategyproofness (SP) An agent will not benefit by misreporting its destination and taking a subsequent ride from that destination to the true one: letting F' being the cost after misreporting of agent i, we require that $p_i(\boldsymbol{c}_{-i}, c'_i; F') + |c_i - c'_i| \geq p_i(\boldsymbol{c}; F)$. Note that $|c_i - c'_i|$, is a lower bound on the cost of travel between i and i'.

Group strategyproofness (GSP) A set of agents will not all benefit by misreporting their destinations and taking a ride from that destination to the real one. Fixing $S \subseteq N$, let $\boldsymbol{c}'_S$ be the costs reported by S (whereas $\boldsymbol{c}_S$ is their true costs), and let F' be the new total cost; then for some $i \in S$: $p_i(\boldsymbol{c}_{-S}, \boldsymbol{c}'_S; F') + |c_i - c'_i| \geq p_i(\boldsymbol{c}; F)$.

Additivity (Add) If a single agent decides to split (e.g., a couple of friends changing their plans and instead of going together to a faraway show, go each to their own home) such that their individual shortest paths sum up to their previous joint one, and without it effecting the overall time of the whole route, the pricing for other agents does not change: For $i \neq j \in N$, $p_i(\boldsymbol{c}; F) = p_i(c_1, \ldots, c_{j-1}, c'_j, c_j - c'_j, c_{j+1}, \ldots, c_n; F)$.

Player Monotonicity (Mono) If $c_i < c_j$, then $p_i(\boldsymbol{c}; F) \leq p_j(\boldsymbol{c}; F)$.

Scale invariance (SI) Scaling the entire instance by a constant $\alpha \in \mathbb{R}_+$ proportionally scales the allocations by α: For all $i \in N$, $p_i(\alpha\boldsymbol{c}; \alpha F) = \alpha p_i(\boldsymbol{c}; F)$

3 Incompatible Axiom Sets

Many of the axioms we are interested in turn out to be incompatible. This section explores different combinations of axioms which are not compatible with efficiency, symmetry, IR, and/or non-negativity.

Theorem 1. *Efficiency, symmetry, non-negativity and group-strategyproofness are incompatible.*

Proof. Consider a set of $n > 6$ agents, each with $c_i = \frac{3}{2}$. Their route from the source has a shared portion with cost $\frac{3}{4}$, and then each agent has a different portion with cost $\frac{3}{4}$. The overall cost of servicing all of them is $F = \frac{3}{2}n$ ($\frac{3}{2}$ to get to the first destination, and $\frac{3}{2}$ between each agent), and thanks to symmetry, this means $p_i = \frac{3}{2}$ for each $i \in N$.

Now, let us take the set $S \subset N$, $|S| = \lfloor \frac{2n}{3} \rfloor + 2$, which all manipulate to announce their destination as to the edge of the shared route (so, for $i \in S$, $c_i = \frac{3}{4}$). In order to prevent successful manipulation, at least one $i \in S$ should have $p_i \geq \frac{3}{4}$, and due to symmetry, they all must have this. The overall cost of the changed set, N' is $F' = \frac{3}{2}|N \setminus S| = \frac{3}{2}(\lceil\frac{n}{3}\rceil - 2)$, This means $\sum_{i \in S} p_i \geq |S|\frac{3}{4} > F'$, and due to non-negativity, $\sum_{i \in N'} p_i > F'$, contradicting efficiency.

Theorem 2. *The only pricing function that is efficient, IR, non-negative and additive is proportional sharing*

$$p_i = F \frac{c_i}{\sum_{j=1}^{n} c_j}$$

Proof. That the suggested function is efficient, IR, non-negative and additive is clear (recall that $F \leq \sum_{j=1}^{n} c_j$, as otherwise, the optimal cost would be for people to take a vehicle to their destination on their own).

Thanks to non-negativity and additivity, we know that an agent with $c_i = 0$ will have $p_i = 0$. First, let us assume that for all $i \in N$, $c_i \in \mathbb{Q}$, i.e., $c_i = \frac{s_i}{t_i}$, $s_i, t_i \in \mathbb{N}$. We divide each agent i into s_i agents with cost of $\frac{1}{t_i}$ each, and thanks to efficiency, symmetry and additivity, we know that for all agents, $p(\frac{1}{\prod_{j=1}^{n} t_i}, \ldots, \frac{1}{\prod_{j=1}^{n} t_i}, F) = \frac{F}{\sum_{k=1}^{n} s_k \prod_{j=1;j\neq k}^{n} t_j}$, since there are $\sum_{k=1}^{n} s_k \prod_{j=1;j\neq k}^{n} t_j$ elements in that sum. Therefore, and thanks to additivity, we know $p_i(c_1, \ldots, c_n, F) = p_i(c_i, \frac{1}{\prod_{j=1}^{n} t_i}, \ldots, \frac{1}{\prod_{j=1}^{n} t_i}, F) = F \frac{s_i \prod_{j=1;j\neq i}^{n} t_i}{\sum_{k=1}^{n} s_k \prod_{j=1;j\neq k}^{n} t_j}$.

Looking closer at the fraction $\frac{s_i \prod_{j=1;j\neq i}^{n} t_i}{\sum_{k=1}^{n} s_k \prod_{j=1;j\neq k}^{n} t_j}$, notice that

$$\frac{\sum_{k=1}^{n} s_k \prod_{j=1;j\neq k}^{n} t_j}{s_i \prod_{j=1;j\neq i}^{n} t_i} = \sum_{k=1}^{n} \frac{s_k t_i}{s_i t_k} = \sum_{k=1}^{n} \frac{c_k}{c_i} = \frac{\sum_{k=1}^{n} c_k}{c_i}$$

Therefore, $p_i(c_1, \ldots, c_n, F) = \frac{c_i}{\sum_{k=1}^{n} c_k} F$.

Now, because any additive function that is continuous at a single point is continuous throughout, and p_i is continuous at 0 (thanks to non-negativity and additivity), this means what we have shown for $c_i \in \mathbb{Q}$ is true for $c_i \in \mathbb{R}$ as well.

Corollary 1. *Efficiency, IR, non-negativity, additivity and strategyproofness are not compatible.*

Proof. Thanks to Theorem 2 we only need to show the pricing function $p_i(c_1, \ldots, c_n, F) = F \frac{c_i}{\sum_{j=1}^{n} c_j}$ is not strategyproof. Taking again the example from Theorem 1, examine a set of $n > 2$ agents, each with $c_i = \frac{3}{2}$, which have a shared route from the source of cost $\frac{3}{4}$, and then each with a different route of length $\frac{3}{4}$. The overall cost of servicing all of them $F = \frac{3}{2}n$, and each agent pays $\frac{3}{2}$. Should one of them choose to deviate to edge of their shared route, it would cost that agent $\frac{3}{2}(n-1)\frac{\frac{3}{4}}{\frac{3}{2}(n-1)+\frac{3}{4}} + \frac{3}{4} < \frac{3}{2}$. Hence, it was beneficial for the agent to deviate.

4 The CEG Mechanism

The constrained equal gains (CEG) mechanism [15, 3] takes an instance $\langle \boldsymbol{c}; F\rangle$, and calculates λ, the solution of

$$\sum_{i\in N} \min\{\lambda, c_i\} = F.$$

Under this mechanism, player i pays $CEG_i = \min\{\lambda, c_i\}$. This means we can essentially divide players into *large* ones if they pay λ, or *small* if they pay their own cost (c_i), which is less than λ. This is equivalent to stating that if there are t large players,

$$\lambda = \frac{1}{t}\left(F - \sum_{j=1}^{n-t} c_j\right).$$

It will be easier to consider an alternative characterization of CEG. Given $\langle \boldsymbol{c}; F\rangle$ we define $Q(i)$ as:

$$Q(i) = \frac{F - \sum_{j<i} c_j}{n-i+1}$$

thus, $Q(1) = \frac{F}{n}$, $Q(2) = \frac{F-c_1}{n-1}$ and so on. CEG can be then defined as follows: we try setting $\lambda = Q(1)$. If there are agents for which it is too high a cost, they pay their own cost, and we try to set $\lambda = Q(2)$, and continue this process until we find a suitable λ.

4.1 CEG – a Strategyproof Mechanism

In this section, we show that CEG is strategyproof, assuming that an agent cannot change the total cost of the ride too much by deviating.

Consider the case where some agent $i \in N$ deviates. Suppose i's true destination is d_i, at distance c_i from s, and the cost of the original ridesharing route is F. Suppose agent i announces their destination to be at d'_i, which is c'_i from the source, making the new route cost F'. Now, we define $\delta = c(d'_i, d_i)$ — the cost for agent i to get to the true destination from the one announced to the ridesharing service. We also define $\delta' = c(d_i, d'_i)$ as the cost of going in the opposite direction. Since we do not assume the graph is necessarily symmetric, it is possible that $\delta \neq \delta'$. Let t be the number of big players before the deviation and t' be the number of big players after the deviation. We provide a sketch of the proof of Theorem 3, which states the strategyproofness of CEG, conditional on a relation between δ, δ' and t.

Theorem 3. *CEG is strategyproof if $\delta' \leq (t-1)\delta$ for every agent on the graph.*

Proof. Consider the original instance $\langle \boldsymbol{c}; F\rangle$, and the revised instance $\langle \boldsymbol{c}_{-i}, c'_i; F'\rangle$ resulting from the deviation described above. We wish to show such a deviation cannot be profitable for agent i.

Case 1: Agent i ends up being a small agent (in F')

Agent i is therefore paying $c_i' + \delta$ (c_i' to get to d_i', and δ to get from there to d_i). Since c_i is the minimal distance from s, $c_i \leq c_i' + \delta$, and since in CEG no agent ever pays more than their cost (so agent i does not pay more than c_i), this cannot result in a profitable deviation for agent i.

Case 2: Agent i ends up being a big agent (in F')

If $F' \geq F$, then the amount paid by a big agent only grows, so if agent i is to profit from deviating then $F' < F$, which means t (number of big agents) changed to t' such that $t' > t$. Also, note that since $F \leq F' + \delta + \delta'$, we know from the theorem's condition that $F \leq F' + t\delta$.

The amount paid by agent i before deviation is $\min(c_i, \frac{F-x}{t})$, for $x = \sum_{j=1}^{n-t-1} c_j$. The amount paid by the agent after deviation is $\frac{F'-x'}{t'}$, for $x' = \sum_{j=1}^{n-t'-1} c_j$. So in order for the deviation to be profitable

$$\delta + \frac{F'-x'}{t'} < \frac{F-x}{t}$$

Since $F - x = F - x' - y$ for $y = \sum_{j=n-t'}^{n-t-1} c_j$, and since Also, thanks to the theorem requirement, we know $F' \geq F - t\delta$, and therefore this can be written as

$$\delta + \frac{F-x'-t\delta}{t'} < \frac{F-x'-y}{t}$$
$$t't\delta + t(F-x') - t^2\delta < t'(F-x') - t'y$$

For agent $c_{n-t'}$, the smallest value turned big by agent i's deviation, we know

$$\frac{F'-x'}{t'} \leq c_{n-t'} < \frac{F-x}{t'}$$

. Combining that $F' \geq F - t\delta$, this means

$$c_{n-t'} \geq \frac{F-x'-t\delta}{t'}$$

Hence

$$y = \sum_{j=n-t'}^{n-t-1} c_j \geq \sum_{j=n-t'}^{n-t-1} c_{n-t'} = (t'-t)c_{n-t'}$$
$$\geq (t'-t)\frac{F-x'-t\delta}{t'}$$

Returning to our previous equation, this means

$$t't\delta + t(F-x') - t^2\delta < t'(F-x') - t'y$$
$$t't\delta + t(F-x') - t^2\delta < t'(F-x') - t'(t'-t)\frac{F-x'-t\delta}{t'}$$
$$t't\delta + t(F-x') - t^2\delta < t'(F-x') - (t'-t)(F-x') + t(t'-t)\delta$$
$$t't\delta - t^2\delta < t(t'-t)\delta$$

which is a contradiction.

If we drop the assumption that $\delta' \leq (t-1)\delta$, then CEG is not strategyproof. Consider Figure 1 with any $\varepsilon > 0$. The true location of agent 1 is at 1, and $c_1 = 11, c_2 = 14, F = 16 + \varepsilon$ so both agents will pay $8 + \frac{\varepsilon}{2}$ under CEG. When agent 1 reports his location as 1^*, $F' = 14$ so agent 1 pays $7 + 1 = 8 < 8 + \frac{\varepsilon}{2}$, so in fact the bound is tight (and, indeed, when $t = 2$, the condition is equivalent to a symmetric graph, though this is not true when t is larger).

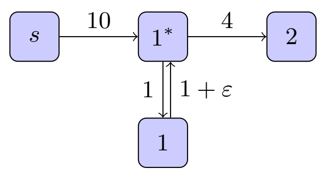

Fig. 1. A possible deviation

4.2 CEG is Uniquely Strategyproof for up to 4 Players

In Section 4.1 we showed that CEG is strategyproof, given some reasonable assumptions on the graph structure. Now, we would like to see if CEG is the *only* strategyproof cost-sharing mechanism. We shall prove that for 4 players or less, if we assume $c_i, F \in \mathbb{Q}$, and want a symmetric, individually rational, player monotone, efficient and scale invariant mechanism, CEG is indeed unique.

Lemma 1. *Suppose a strategyproof mechanism charges a passenger c_i in the setting $\langle \boldsymbol{c}; F \rangle$; then for any $c_i^* \leq c_i$, the mechanism must charge the passenger c_i^* in any setting $\langle \boldsymbol{c}_{-i}, c_i^*; F \rangle$.*

Proof. Consider a setting generated by the graph described in Figure 2. The mechanism cannot charge i more than c_i^* when reporting destination as i^* (due to IR). If it charges i less than c_i^* for reporting destination i^*, then i can profitably deviate to i^*, be charged a sum of $p_i^* < c_i^*$, and travel the rest of the way alone for a cost of $c_i - c_i^*$. This results in a total payment of strictly less than c_i, contradicting strategyproofness.

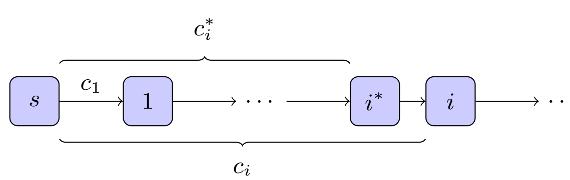

Fig. 2. A possible deviation if the condition in Lemma 1 does not hold.

Theorem 4. *For a 4 player setting where $c_1 \leq c_2 \leq c_3 \leq c_4 \leq F$ with rational costs, when Theorem 3's conditions are satisfied for any possible deviation by players, CEG is the unique SP, IR, symmetric, player monotone, efficient and scale invariant fare division.*

Proof. For this section, we consider the case where c_1, c_2, c_3, c_4, F are all rational. Given scale invariance, it suffices to consider settings with integer c_1, c_2, c_3, c_4, F. Furthermore, by scale invariance, in instances where $c_1 \leq c_2 \leq c_3 \leq c_4 < F$, we can assume that $c_4 < F - 6$ (by multiplying the whole instance by a suitably large constant). By strategyproofness, p_i needs to be continuous both in c_i and F, therefore we can extend to cases $c_1 \leq c_2 \leq c_3 \leq c_4 \leq F$.

We proceed to prove the theorem by induction on $c_1+c_2+c_3+c_4+F$. When $c_1 = c_2 = c_3 = c_4 = F = 0$, each player pays 0 by IR, so equal sharing is used. Suppose CEG is used whenever $c_1 + c_2 + c_3 + c_4 + F < k$. Consider an instance with $c_1 + c_2 + c_3 + c_4 + F = k$. For the next parts, we check the cases based on the number of big players in the system, as well as the number of big players remaining after the selected player attempts to make a deviation. Below, Case 1 is the case where there are 4 big players. The rest 3 cases, omitted due to space constraints have 3, 2, and a single big player respectively. We make use of Lemma 1 to transition between cases.

Case 1: Consider the case where $\frac{F}{4} \leq c_1 \leq c_2 \leq c_3 \leq c_4 < F - 6$

We want to show that equal sharing must be used. In the scenarios we consider below, the original locations of players 1,2,3 and 4 are at nodes 1,2,3, and 4 respectively, and the deviating player moves from node p to p^*.

Case 1a: If $c_1 = c_2 = c_3 = c_4$, by symmetry equal sharing must be used.

Case 1b: If $c_1 < c_2 = c_3 = c_4$, consider Figure 3, and set $a = F - c_4 - 6$ (note $F - c_4 - 6 \geq 0$ since we assume $c_4 \leq F - 6$). The best possible route is $s \to 4 \to 3 \to 2 \to 1$, with a cost of F, or $s \to 4 \to 3 \to 2^* \to 2 \to 2^* \to 1$, which also costs F. Now suppose player 2 reports his destination as 2^* instead of 2. The best route is now $s \to 4 \to 3 \to 2^* \to 1$, with a cost of $F - 4$. This is now an instance of the form $\langle c_1, c_2 - 1, c_3, c_4; F - 4\rangle$, and by the induction hypothesis equal sharing must be used. Therefore, each player pays $\frac{F}{4} - 1$, and so player 2 pays $\frac{F}{4}$ in total to his destination. Therefore, player 2 cannot pay more than $\frac{F}{4}$ originally, otherwise this would be a viable manipulation, contradicting strategyproofness. Since $c_2 = c_3 = c_4$, by symmetry players 3 and 4 also cannot pay more than $\frac{F}{4}$, and so by efficiency equal sharing must be used in the original scenario.

Case 1c: If $c_1 \leq c_2 < c_3 = c_4$, consider the graph in Figure 4, and set $a = F - c_4 - 6$. A possibility for the optimal route is $s \to 4 \to 3^* \to 3 \to 2 \to 1$, costing F. Now suppose player 3 reports his destination as 3^* instead of 3. The best route is now $s \to 4 \to 3^* \to 2 \to 1$, which costs $F - 4$. With the same argument as in case 1a, player 3 cannot pay more than $\frac{F}{4}$ in the original case.

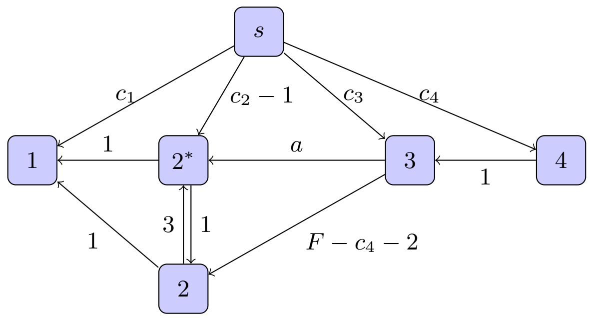

Fig. 3. Graph for case 1b

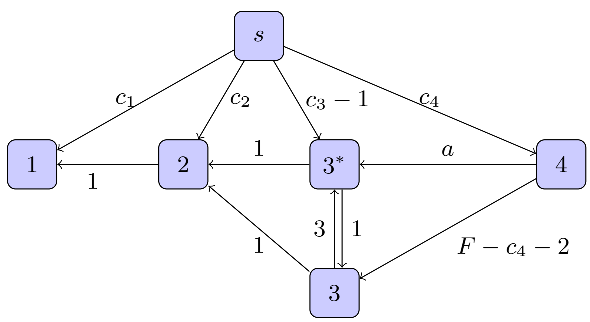

Fig. 4. Graph for case 1c

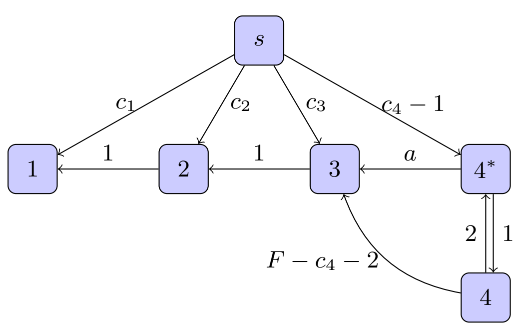

Fig. 5. Graph for case 1d

By symmetry, player 4 also pays $\frac{F}{4}$. By player monotonicity, players 1 and 2 cannot pay more than $\frac{F}{4}$, so by efficiency equal sharing must be used in the original setting.

Case 1d: If $c_1 \leq c_2 \leq c_3 < c_4$, consider the graph in Figure 5, and set $a = F - c_4 - 5$. The original best possible route is $s \to 4^* \to 4 \to 3 \to 1$, costing F. If player 4 reports his destination as 4^* instead of 4, the best possible route is now $s \to 4^* \to 3 \to 2 \to 1$, costing $F - 4$. With the same argument as in case 1a, player 3 cannot pay more than $\frac{F}{4}$ in the original case. By player monotonicity, players 1, 2 and 3 each cannot pay more than $\frac{F}{4}$, so by efficiency equal sharing must be used in the original setting.

The rest of the cases follow similar arguments; the appropriate graphs can be constructed following the same structure as above by varying the value for a. They are omitted due to space constraints. □

Theorem 5. *For a 2 (or 3) player setting where $c_1 \leq c_2 (\leq c_3) \leq F$ with rational costs, and when Theorem 3's conditions are satisfied for the deviating player, CEG is the unique SP, IR, symmetric, player monotone, efficient and scale invariant fare division.*

Proof (Sketch of proof). Set c_1 and c_2 to 0 in the proof of $n = 4$.

5 Simulations

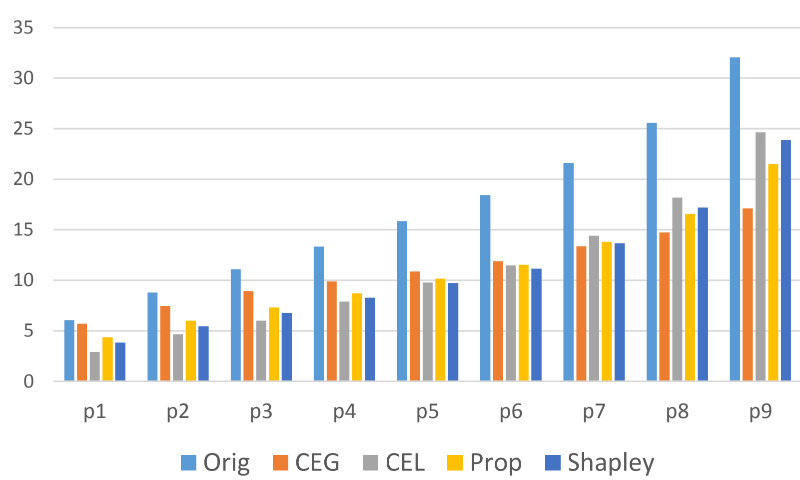

Fig. 6. Light blue bar at the left of each passenger indicates the average $\frac{c_i}{F}$; other bars indicate average cost by sharing rule (normalized by F).

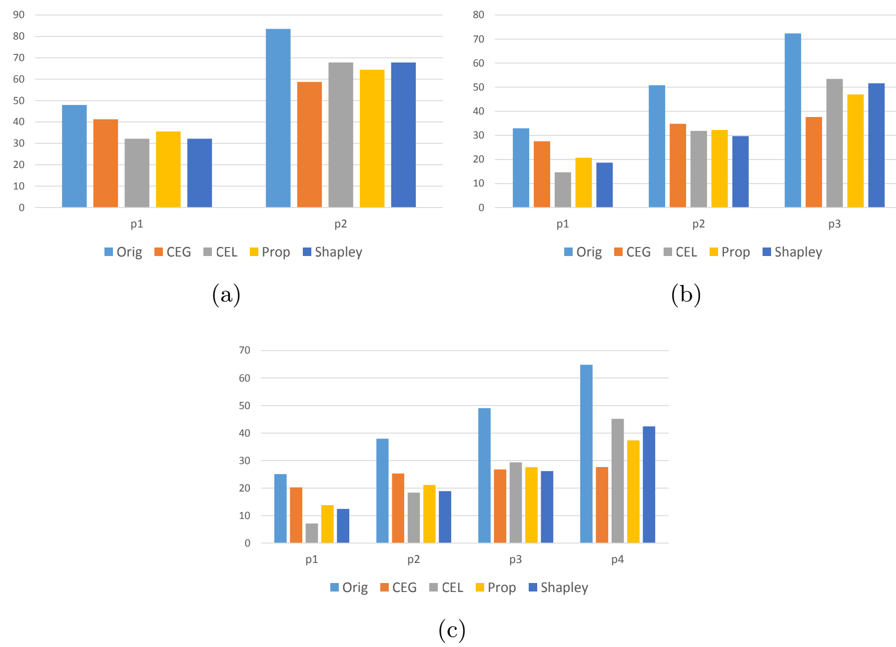

Fig. 7. Average percentage payment allocations with 2-4 agents under each allocation rule.

We examine our cost sharing methods in practice by looking at real taxi rides[7] made from the same (approximate) location, i.e. within a radius of 300 meters. We ran a total of 423 simulations, for a total of 1416 shared rides (see Table 2 for details) The costs were measured in seconds to reach the destination at 17:00, calculated using the Google Maps Distance Matrix API; the optimal set of routes was calculated using dynamic programming. We performed this routine by optimally allocating sets of 9 players trying to get to their destination from the same place at the same time, which we divided to separate taxis, limiting cab capacity to 4.

[7] `http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml`

Rule	**Min**	**Max**	**Stdev**
CEG	0.076	16.726	5.623
CEL	1.287	8.969	2.926
Prop	0.812	12.238	3.701
Shapley	0.800	11.418	3.586

Table 1. Measures on utility properties of different sharing rules, averaged over 423 simulations of 9 players

# of passengers in cab	# of instances
1	275
2	346
3	340
4	455

Table 2. Number of instances generated, by number of passengers in ride

We compare the following cost-sharing methods:

Proportional: Players pay in proportion to their cost: $p_i = \frac{c_i}{\sum_i c_i} F$.
Shapley: Players pay the Shapley value (this is the method used by Spliddit); in more detail, for every subset of players $S \subseteq N$, let $F(S)$ be the fare that S would have paid if they were the only players to be dropped off. Given $S \subseteq N \setminus \{i\}$, let $m_i(S)$ be $F(S \cup \{i\}) - F(S)$, i.e. the marginal contribution of i to S; the Shapley value of i is $\frac{1}{n!} \sum_{S \subseteq N \setminus \{i\}} |S|!(n - |S| - 1)! m_i(S)$ (see [5] for details).
CEL: The savings from the shared ride are equally divided between passengers. More formally, each player pays $\max\{0, c_i - \mu\}$, where μ is the solution to $F = \sum_{i=1}^{n} \max\{0, c_i - \mu\}$.
CEG: Defined above in Section 4.

Of the four measures, only the Shapley value uses information beyond $\langle \boldsymbol{c}; F \rangle$, as it computes ride costs for every subset S of players. This means that the Shapley value is able to leverage more of the underlying graph structure than the other allocation methods, and is, on the other hard, much more demanding computationally.

Figure 6 describes the average percentage of the overall fare paid by the players in each instance. While $c_1 \leq c_2 \leq \cdots \leq c_9$ in all instances, players were divided into taxis for an optimal driving time. Under CEG, players 1,2 and 3 pay fares close to their costs, while players $4, \ldots, 9$ pay significantly less than their cost. This translates to closer equality of payment, and higher social disparity in the division of the ridesharing surplus: lower cost players gain next to nothing by participating in the shared ride, whereas higher cost players are far better off. This can be observed in Table 1: the average maximal normalized utility (measured as $\max_{i \in N} \frac{c_i - p_i}{F}$) is higher for CEG than it is for any other fare allocation method.

In 46.5% of multiplayer instances in our data, all passengers in a cab had $c_i \geq \frac{F}{n}$, i.e., all passengers were big. In this case, CEG simply allocates a fare of $\frac{F}{n}$ to all players. Thus, to some extent, CEG uses very little information about players' costs, as opposed to more subtle mechanisms such as proportional or the Shapley fare allocation. CEL does the opposite of CEG: smaller cost players pay almost nothing, and *savings* are divided very evenly. This is easily seen in Table 1, as well as Figure 7: the average maximal normalized utility of any player is significantly lower for CEL than it is for other fare division methods. As can be seen in Figures 5 and 5, the majority of losses and gains provided by CEG and

CEL affect the players with the lowest and highest costs. The Shapley value is the only allocation which accounts for players' location on the graph; as a result, the Shapley value allocations show much more local variation than the other methods. However, despite its complex structure, Shapley-based allocations are quite close those outputted by the proportional method (see Figure 7).

6 Conclusions and Future Work

This paper examines the cost-sharing problem in the context of ridesharing. In doing so, we presented a set of desirable properties; after showing several impossibility theorems regarding their combinations, showed that CEG is the unique strategyproof mechanism for up to 4 participants – the number that can fit in common cabs. Despite CEG's attractive properties, our simulations show it to be very rigid – CEG is strategyproof because it strives to make players' payments as equal as possible. This benefits mainly players with high costs over players with small ones, who see little benefit from ridesharing. In some sense, our results show that no reasonable strategyproof mechanism can ensure that all players strictly benefit from sharing the ride, unless the ridesharing savings are very significant compared to *each* of their individual cost.

Such fare division method is, therefore, not very practical in the real world – many participants will see very little benefit in ridesharing using it. Thus, non-strategyproof techniques will inevitably be used, with a potential loss of revenue. A possible direction for further research will try and examine ϵ-strategyproofness or other truth approximation methods, to try and guarantee that manipulation will not be very worthwhile. Indeed, while we show that CEG is uniquely strategyproof for the case of up to four players, we do not have a general proof for n players; this would certainly be an important first step in understanding the general structure of strategyproof cost sharing.

A few other cost sharing methods for this problem were examined in our simulations; these methods are also axiomatically justified (the axiomatic treatment of the Shapley value [16] is perhaps the most well-known, and [15] takes that approach to cost sharing, but using properties relevant to the ridesharing setting is desirable). This means that one can go about choosing a cost sharing mechanism in a rather principled way, choosing the axioms that are most pertinent to the ridesharing application. Naturally, there is also plenty to investigate in the real-world data that is becoming increasingly available for these topics. For example, exploring people's distribution when taking rides may allow for algorithms that are tailor-made for common distributions, allowing various properties which will apply only in specific settings.

Acknowledgements

This work was partially supported by NSERC grant 482671, by MOE Grant no. R-252-000-625-133, and by NRF Fellowship no. R-252-000-750-733.

References

1. Agatz, N., Erera, A., Savelsbergh, M., Wang, X.: Optimization for dynamic ride-sharing: A review. European Journal of Operational Research 223(2), 295–303 (2012)
2. Agatz, N., Erera, A., Savelsbergh, M., Wang, X.: Dynamic ride-sharing: A simulation study in metro Atlanta. Transportation Research Part B: Methodological 45(9), 1450–1464 (2011)
3. Aumann, R.J., Maschler, M.: Game theoretic analysis of a bankruptcy problem from the talmud. Journal of economic theory 36(2), 195–213 (1985)
4. Bistaffa, F., Farinelli, A., Chalkiadakis, G., Ramchurn, S.: Recommending fair payments for large-scale social ridesharing. In: Proceedings of the 9th ACM Conference on Recommender Systems. pp. 139–146. RecSys'15 (2015)
5. Chalkiadakis, G., Elkind, E., Wooldridge, M.: Computational Aspects of Cooperative Game Theory. Synthesis Lectures on Artificial Intelligence and Machine Learning, Morgan and Claypool (October 2011)
6. Frisk, M., Göthe-Lundgren, M., Jörnsten, K., Rönnqvist, M.: Cost allocation in collaborative forest transportation. European Journal of Operational Research 205(2), 448–458 (2010)
7. Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M., Wang, X., Koenig, S.: Ridesharing: The state-of-the-art and future directions. Transportation Research Part B: Methodological 57, 28–46 (2013)
8. Galland, S., Knapen, L., Yasar, A., Gaud, N., Janssens, D., Lamotte, O., Koukam, A., Wets, G.: Multi-agent simulation of individual mobility behavior in carpooling. Transportation Research Part C: Emerging Technologies 45, 83–98 (2014), advances in Computing and Communications and their Impact on Transportation Science and Technologies
9. Goldman, J., Procaccia, A.: Spliddit: Unleashing fair division algorithms. SIGecom Exchanges 13(2), 41–46 (2015)
10. Johari, R.: The price of anarchy and the design of scalable resource allocation mechanisms. In: Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V. (eds.) Algorithmic game theory. Cambridge University Press (2007)
11. Kamar, E., Horvitz, E.: Collaboration and shared plans in the open world: Studies of ridesharing. In: Proceedings of the 21st International Joint Conference on Artifical Intelligence. pp. 187–194. IJCAI'09 (2009)
12. Kleiner, A., Nebel, B., Ziparo, V.: A mechanism for dynamic ride sharing based on parallel auctions. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence. pp. 266–272. IJCAI'11 (2011)
13. Li, B., Krushinsky, D., Reijers, H.A., Van-Woensel, T.: The share-a-ride problem: People and parcels sharing taxis. European Journal of Operational Research 238(1), 31–40 (2014)
14. Ma, S., Zheng, Y., Wolfson, O.: Real-time city-scale taxi ridesharing. IEEE Transactions on Knowledge and Data Engineering 27(7), 1782–1795 (2015)
15. Moulin, H.: Axiomatic cost and surplus sharing. In: Arrow, K.J., Sen, A.K., Suzumura, K. (eds.) Handbook of Social Choice and Welfare, vol. 1, chap. 6, pp. 289–357. Elsevier (2002)
16. Shapley, L.S.: A value for n-person games. Contributions to the Theory of Games 2(28), 307–317 (1953)
17. Zhao, D., Zhang, D., Gerding, E.H., Sakurai, Y., Yokoo, M.: Incentives in ridesharing with deficit control. In: Proceedings of the 13th International Joint Conference

on Autonomous Agents and Multiagent Systems (AAMAS). pp. 1021–1028. Paris, France (May 2014)