RICE UNIVERSITY

Essays in Mechanism Design and Decision Theory

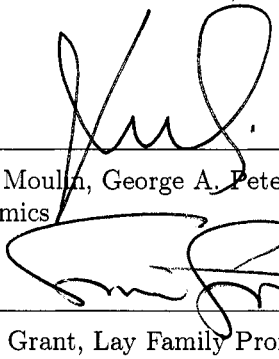
by

Hatice Özsoy

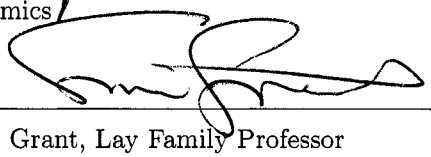
A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

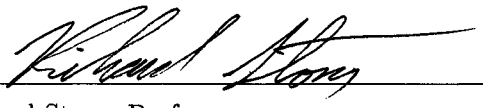
APPROVED, THESIS COMMITTEE



Hervé Moulin, George A. Peterkin Professor, Chair
Economics



Simon Grant, Lay Family Professor
Economics



Richard Stong, Professor
Mathematics

HOUSTON, TEXAS
MAY, 2007

UMI Number: 3309932

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3309932

Copyright 2008 by ProQuest LLC.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 E. Eisenhower Parkway
PO Box 1346
Ann Arbor, MI 48106-1346

Abstract

Many economic decisions rely on information that is privately owned by the agents, who may have incentives to misreport this information. The strategic possibilities of the agents may take different forms, depending on the economic environment. The first chapter studies the scheduling problem, and considers coalitional maneuvers in the form of splitting and transferring jobs. The second chapter explores strategic behavior in the form of merging by users in the minimum cost spanning tree problem. The third chapter is about decision making under uncertainty, where we give an alternative characterization of individual preferences that are based on beliefs.

Acknowledgements

The first two chapters of this dissertation were written under the supervision of Hervè Moulin. I am indebted to him for his valuable guidance over the last five years. These chapters also benefited from the comments of Richard Stong. The third chapter is joint work with Simon Grant and Ben Polak. I am thankful to Dr. Grant for encouraging me to participate in this project, and for very helpful discussions. I am grateful to Gil Eris for always developing elaborate questions that helped me give better answers.

Table of Contents

Abstract		ii
Acknowledgements		iii
Introduction		1
Chapter 1. Coordinated splitting in probabilistic scheduling		6
1.1	Introduction	6
1.2	Related literature	8
1.3	The model	9
1.4	Normative properties of scheduling rules	11
1.5	Coordinated split-proofness	15
1.6	Transfer-proofness	21
1.7	Concluding remarks	24
Chapter 2. A characterization of Bird's rule		25
2.1	Introduction	25
2.1.1	An illustrative example	25
2.1.2	Overview of main results	28
2.2	Preliminaries	29
2.2.1	Jarník-Prim-Dijkstra algorithm and related solutions	29
2.3	Merge-Proofness	31
2.4	Discussion	40
2.4.1	Implications of the main result	40
2.4.2	Independence of the axioms	41

Table of Contents (Cont'd)

Chapter 3. Purely subjective probabilistic sophistication with Stochastic Monotonicity 44

3.1 Introduction 44

3.2 Preliminaries 46

3.2.1 Setup 46

3.2.2 Subjective expected utility 47

3.3 Probabilities via Stochastic Monotonicity 50

References 54

Appendix: Proofs 59

Lemma 1 59

Theorem 2 60

Proposition 1 64

Introduction

Economic decisions in many environments rely on information that is privately owned by the agents. In many situations, the agents have incentives to misreport this private information. For example, in an exchange economy, one type of private information is an individual's preferences over consumption bundles. Under the competitive allocation for an exchange economy, an agent may achieve a higher utility by pretending to have different preferences than the actual ones. However, misrepresenting one's preferences is not the only way to manipulate an allocation rule. Postlewaite (1979) studies several different manipulations in the context of a pure exchange economy. An agent may destroy or hide part of his endowment, in order to achieve a higher utility. Such maneuvers may involve coalitions of agents instead of a single agent. A group of agents can reallocate their endowments within the coalition so as to achieve a higher utility for all members of the coalition.

Depending on the economic environment, the strategic possibilities of the agents may take very different forms. The first two chapters in this dissertation study strategic behavior in two different environments. The first chapter looks at the *scheduling problem.* Imagine a set of users arriving at the same time to a server to process their jobs. Assume that all users have the same linear cost of waiting, i.e. that each user incurs a cost of x if his waiting time is x units. A *scheduling rule* assigns a probability to each possible ordering of the jobs. Suppose the set of users is given by $N = \{1, 2, .., n\}$. The *Uniform rule* is the rule that assigns the probability $\frac{1}{n!}$ to each ordering of the jobs. Given that the scheduling rule is probabilistic, the users face random waiting times. We assume that the users are risk neutral, so that they care only about the expected wait.

We assume that the server cannot monitor the identity of the users. If new identities are easy to create, then a user could split his job into several components and submit each component under different aliases. He would have an incentive to do this if he can decrease his expected wait by doing so. Moulin (2004b) studies this issue, and identifies a large class of scheduling rules that are immune to splitting by a single user. Here, we study the extension of splitting by a single user to splitting by coalitions, and find that the results in Moulin (2004b) are not robust to such an extension. Say a scheduling rule is *coordinated split-proof* if it is not immune to coalitional splitting

maneuvers. We find that, if we restrict attention to scheduling rules that satisfy *anonymity* and *separability* (a tractability condition), all rules that are coordinated split-proof are essentially the same as the Uniform rule.

A transfer maneuver is another form of strategic behavior in this context. It is feasible in a wider range of environments than splitting, since it does not rely on the creation of new aliases. Let the job profile be given by $(x_1, x_2, ..., x_n)$. Given a nonnegative function w, Moulin (2004b) introduces the $w-$*quasi- proportional scheduling rule* in the following way: first assign the job that is going to be served last, where the probability that x_n is served last is given by $\frac{w(x_n)}{\sum_{j=1}^{n} w(x_j)}$; then choose in the same way the job that will be served next to last, where the probability that x_{n-1} is served last is given by $\frac{w(x_{n-1})}{\sum_{j=1}^{n-1} w(x_j)}$; and so on. We find that a $w-$quasi-proportional rule is immune to pairwise transfers if w is concave and nondecreasing. If we require that a $w-$quasi-proportional rule satisfy *scale invariance*, e.g. if we double everyone's job size, we should double everyone's expected wait, then we know that w is of the form $w(x) = x^{\alpha}$ for some $\alpha \geq 0$. The Uniform rule, for instance, corresponds to $\alpha = 0$. We call the set of such $w-$quasi-proportional rules the *power family*. If we allow for transfers between three users, the only rule from the power family that is immune to transfer maneuvers is the Uniform rule.

The second chapter of the dissertation studies strategic behavior in the minimum cost spanning tree problem. There is a set of towns each of which need to connect to the source of some service (like a power plant) directly or indirectly through other towns. Let 0 denote the source and N the set of users. An input in this problem is a cost matrix $[c_{ij}]_{ij \subset N_0}$ that specifies the cost of direct connection between any $i, j \in N \cup 0$. The users are endowed with quasi-linear preferences in money and monetary transfers are feasible, so that efficiency requires that we construct the cheapest possible connected network, i.e. a minimum cost spanning tree (mcst). We are interested in sharing the cost of the mcst among the users.

Several *cost allocation rules* have been proposed for mcst problems. The first such rule is *Bird's rule* suggested by Bird (1976). Suppose that there is a unique mcst and, for each $i \in N$, let $\alpha(i)$ denote the immediate predecessor of i in the mcst. In other words, $\alpha(i)$ is directly connected to i

in the mcst, and lies on the path from 0 to i. Bird's rule charges user i the cost of the link $(\alpha(i)\ i)$.

Claus and Kleitman (1973) argued that immunity to system break-up and mergers by users are desirable properties of cost allocation rules. A standard assumption in the literature since their work is that each coalition of users $S \subset N$ can break away from the system and construct their own network to get the service from the source. Following Bird (1976), we derive a cooperative game from a mcst problem by defining the worth of S to be the cost of a mcst on $S \cup 0$. Now, solutions like *Shapley value* and *nucleolus* find counterparts in the mcst problem. We say that a cost allocation is in the *core* of the derived cooperative game if no group of users can break away from the system in order to decrease their joint costs. Accordingly, we say that a cost allocation rule satisfies *core selection*, if it always picks an allocation in the core Granot and Huberman (1981) showed that mcst games always have a nonempty core. However, they do not have to be *concave*, and the Shapley value may select allocations outside of the core, i.e. it is vulnerable to system break-up. We also criticize the Shapley value and several other rules proposed in the literature on the grounds that they are vulnerable to merging maneuvers by groups of users.

We consider two types of merging as forms of strategic behavior: covert-merging and overt-merging. Under *a covert-merging maneuver*, a coalition $S \subset N$ chooses a user $i^* \in S$ to stay in the problem. From the point of view of the service provider, it is as if the users in $S \backslash i^*$ disappear from the problem. These users will get the service only through user i^* via a private network, the costs of which are borne by S. We assume that S constructs the cheapest private network, i.e. a mcst on S. Hence, when S merges, it has to pay the cost allocation of the merged user i^* in the new problem, plus the cost of the mcst on S. We say that a cost allocation rule is *covert-merge–proof* if there is no coalition that can decrease its total cost through a covert-merging maneuver. We find that there is no covert-merge–proof way of allocating costs.

Under an *overt-merging maneuver*, the private network built by S is available for use when we construct the mcst, and the service provider knows that coalition S merged into i^*. Hence, when it wishes to connect i^* to some $j \notin S$, it can use any of the links (ij) where $i \in S$, and the cost of direct connection between i^* and j in the merged problem is $\min_{i \in S} c_{ij}$. Bird's rule is immune to overt-

merging maneuvers if the original problem has a unique mcst. If we allow for several mcst, then there is no overt-merge–proof cost allocation rule. Our main result in this chapter is an axiomatic characterization of Bird's rule based on its overt-merge–proofness property. On a restricted domain of cost matrices that ensures a unique mcst, Bird's rule is characterized by *overt-merge–proofness*, *core selection*, and *tree invariance*. Tree invariance asserts that two problems with the same tree should have the same cost allocation. We use our characterization result to show that among a desirable family of cost allocation rules introduces in Feltkamp, Tijs, and Muto (1994), the only one that is overt-merge–proof is Bird's rule.

The third chapter of my dissertation is joint work with Simon Grant and Ben Polak. The subject here is decision making under uncertainty, when uncertainty is not described by exogenously given (objective) probabilities. In the presence of objective probabilities, the objects of choice are lotteries, functions that assign a probability to each outcome. Von Neumann and Morgenstern (1944) provided conditions on the choice behavior of the decision maker under which there exists a utility function over outcomes, and when faced with a choice between two lotteries, the decision maker picks the one with higher *expected utility*. Here, the expected utility of a lottery is the sum of the utilities of each outcome, calibrated by the probability of the event that the lottery yields that specific outcome.

In many real life situations involving uncertainty, individuals have to carry out decisions without knowing the probabilities of the states of nature that describe the uncertainty. A commonly used example in the literature since Anscombe and Aumann (1963) is a bet on a horse race, or a horse race lottery. By observing your choices over horse race lotteries, what can we say about your beliefs as to how likely you find each horse to win?

In his seminal work of 1954, Savage modeled situations involving subjective uncertainty in the following way. A set of states S describe the uncertainty, and the outcomes are members of a given set X. An act, e.g. a horse race lottery, is a function from S to X. Savage (1954) imposed seven conditions on choice behavior under which there exists a probability measure π over events and a utility function over the set of outcomes such that, when comparing two acts, she chooses the one

with higher expected utility, when the utility calibrations are done using π. The derived probability π describes the individual's beliefs about the likelihood of events, if we assume that her evaluation of an outcome is independent of the state $s \in S$ she is in. This is a strong assumption that is not satisfied in many contexts. For example, if you are betting on the outcome of the World Cup, you may enjoy \$100 if your favorite team wins more than you would in the event that it loses.

The Allais paradox (1953) demonstrated that most people are not expected utility maximizers, and in particular that they do not conform to Savage's second postulate, the Sure-Thing Principle. Partially motivated by this, Machina and Schmeidler (1992) provided conditions on choice behavior in subjective settings under which the decision maker *acts as if* she is assigning additive probabilities to events, but is not necessarily an expected utility maximizer. They call such an individual as being *probabilistically sophisticated.* Working in a Savage setting, Machina and Schmeidler (1992) characterize probabilistically sophisticated choice behavior by imposing the Savage conditions except his axiom **P2** and strengthening his axiom **P4**. We show that this stronger axiom embodies a weaker version of **P2**, which we call the *Two-Outcome Sure-Thing Principle.* We provide a set of conditions involving $\mathbf{P2}^{**}$ that are equivalent to Machina and Schmeidler's, and this enables us to give an alternative characterization of probabilistically sophisticated choice behavior.

1 Coordinated Splitting in Probabilistic Scheduling[1]

1.1 Introduction

A set of users arrive at the same time to a single server which can process one job at a time. Partially completed jobs are useless, so efficiency requires to serve each job as a whole. Risk neutral users only care about the total wait time until the completion of their own job. If the server cannot monitor the identity of users, scheduling is vulnerable to splitting and transferring maneuvers. Splitting happens when a single user divides her demand into several parts and reports each of these "split" demands under a different alias. The simplest example of a scheduling mechanism that is vulnerable to both of these maneuvers is the *Shortest Job First* mechanism which serves the jobs in increasing order of their size, with a tie-breaking rule.

Splitting by a single user in probabilistic scheduling has been studied by Moulin (2004b). He finds a wide class of scheduling rules that are invulnerable to splitting. Here, we consider the extension of splitting maneuvers to coalitions, which we call *coordinated splitting*, and ask whether the results on splitting by a single user are robust to such an extension. We find that they are not.

Given a group of users S, a coordinated splitting maneuver happens in the following manner. Each user in S splits her job in several parts. The server will allocate slots to serve these "split" jobs. The users in S make an agreement to set a (possibly randomized) priority on the usage of these slots. We say a scheduling discipline satisfies *coordinated split-proofness* if it is not vulnerable to such maneuvers. We find that this property is very restrictive: among scheduling disciplines that are *separable* and *demand monotonic*, all those that satisfy coordinated split-proofness are essentially the same in that the wait time of a user is the same under each of these disciplines (see Theorem 1).

Splitting maneuvers are feasible in various contexts. Networks that operate on the Internet are of specific relevance as different aliases are very easy to create. Examples include mail servers

[1]**An earlier version of the material in this chapter has been submitted to International Journal of Game Theory. I would like to thank two anonymous referees, an associate editor, and William Thomson for helpful comments, which improved the content and presentation.**

and Peer-to-Peer systems. According to Douceur (2002), as the latter usually rely on the users having distinct identities, the creation of new aliases makes them vulnerable to outside threats. The server can sometimes track the identity of users through their IP addresses. However, if associating a job with its beneficiary is not possible, a user could split her job and submit it through other computers (with or without their consent). This is usually very easy for the user as it can be achieved by installing a cheater program in one's computer.[2] Alternatively, the server might be obliged to keep anonymity, perhaps because the jobs involve confidential information about the users. Examples include Freenet[3], or networks where the users share a financial or medical database. See the discussion in Moulin (2004b).

We also look at transfer maneuvers, which concern a group of users S who might find it in their interest to reallocate the total demand of S among themselves. In this respect, our *transfer-proofness* axiom is very close in spirit to the *no advantageous reallocation* axiom discussed in Moulin (1985) and Chun (1988). In our context, after reallocating their demand among themselves, the users in S would be free to choose how to use the slots assigned to S by the server. In other words, they have to decide who will use the first slot available for S, who will be served next, etc. Theorem 2 below specifies a class of scheduling rules that are invulnerable to transferring of jobs between two users. However, if we confine our attention to the subclass of *quasi-proportional scheduling rules* (Moulin (2004b)) and allow for transfers between three or more users, all transfer-proof scheduling rules are again essentially the same as the Uniform rule (see Proposition 1).

A scheduling problem specifies the set of users N and a nonnegative vector of job lengths x. A scheduling rule chooses a randomized ordering of the users in N. Let $|N| = n$. The Uniform rule, for instance, is the scheduling rule that chooses each ordering of the n jobs with equal probability $1/n!$. Another key scheduling rule is the Proportional rule: it first chooses the job to be scheduled last, where the probability of job i being served last is $\frac{x_i}{x_1+\cdots+x_n}$; it then chooses in the same way the job scheduled next to last among the remaining jobs, and so on. A scheduling method specifies

[2] In the case of KaZaA, a Peer-to-Peer system with more that 350 million users, easy to install Cheater and SpeedUp programs provide opportunities to constantly cheat the server.

[3] Freenet is a Peer-to-Peer system that lets users publish and obtain information through their network. They focus on the importance of freedom of speech, a requirement of which is the anonymity of the users.

the expected waits for the n jobs. For instance, the Uniform rule implements the Uniform method which, to each user i, assigns the expected wait $y_i^{Uni} = x_i + \frac{1}{2}\sum_{j\in N\setminus i} x_j$.

The Proportional rule, although invulnerable to splitting by a single user, is vulnerable to coordinated splitting maneuvers. Consider the scheduling problem where the set of users is A, B, and C, with corresponding jobs of size $2, 2, 1$, respectively. Now suppose users A and B make a coordinated splitting maneuver of the following form: they create two aliases A' and B' and report a job of size 1 under each of the aliases A, A', B, B'; they decide who goes first with $1/2$ probability, and whoever goes first uses the first two 1-unit slots available to users A and B. By making such an agreement, each of the users A and B can decrease their expected wait by more than 4%, see Example 1 below. The Uniform rule, on the other hand, is not vulnerable to coordinated splitting maneuvers. The proof is very simple. Suppose a coalition of users S split their jobs. For a user i outside S, the sum $\sum_{j\in N\setminus i} x_j$ is unaffected, so his expected wait is the same after the splitting maneuver. As the Uniform rule yields an efficient vector of expected waits, it follows that the users in S cannot all benefit from the coordinated splitting move.

Section 2 discusses the related literature, and Section 3 introduces the model. In Section 4 we discuss some normative properties of scheduling rules. Theorem 1 in Section 5 below essentially characterizes the Uniform rule as the only rule immune to coordinated splitting moves. A second property of the Uniform rule is its invulnerability to transfer maneuvers, which we study in Section 6.

1.2 Related Literature

Moulin (2004b) notes that the first discussion of splitting and merging maneuvers are found in the rationing problem (Banker (1981), Moulin (1987), De Frutos (1999), Ju (2003)). In that literature, the proportional rule stands out as the only rule immune to such maneuvers.

The scheduling and queuing literature has mostly focused its attention to contexts where monetary transfers are feasible. Under the assumption of quasi-linear preferences, efficiency requires to use the Shortest Job First rule. Then, monetary compensations are used to restore fairness and

incentive compatibility properties (Curiel et. al. (1989), (2002), Klijn and Sanchez (2002), Maniquet (2003), Chun (2004, 2006)).

In queuing problems with money, when the waiting cost of a user is private information, the truthful revelation of these costs has been widely studied. Suijs (1996) proved that, assuming linear costs, it is possible to design a budget-balanced Vickrey-Clarke-Groves mechanism where it is a dominant strategy to reveal true waiting costs. Mitra and Sen (1998) propose conditions on the cost structure under which such mechanisms exist (see also Mitra (2001, 2002)). In the context where job sizes of the users are private information, truthful revelation of job sizes is achieved in Hain and Mitra (2001) using a generalized Vickrey-Clarke-Groves mechanism.

This chapter of my dissertation is inspired by Moulin (2004b). He studies splitting and merging maneuvers in a probabilistic setting. There, the Proportional rule stands out of split-proofness combined with *ranking*[4] or the reduction of total waiting costs. Moulin (2004a) studies splitting, merging and transferring maneuvers in a context where monetary transfers are feasible. His results concerning transfer-proofness are similar to the implications of Theorem 2 and Proposition 1 below: There is a one-dimensional set of scheduling mechanisms that satisfy pairwise transfer-proofness, but there is no "continuous" mechanism that is invulnerable to transfers among three or more users.

1.3 The Model

Most of the following notation is adopted from Moulin (2004b). $\mathcal{N}$ is the set of potential users. Given a finite subset N of $\mathcal{N}$, a *scheduling problem* is a pair (N, x), where $x \in \mathbb{R}_+^N$ is the profile of job sizes belonging to the users in N.[5] We call a user with a zero job size a *null user*, and a user with a positive job a *nonnull user*. Null users are indifferent as to when they leave the queue. Think of a user in a P2P system who is "signed in", but is not actively doing anything. For nonnull users, we interpret the (expected) wait time (or expected delay) of this user, until the completion of her job, as the "disutility" she incurs, and assume that partially completed jobs are useless. The users are

[4] A scheduling method meets ranking if $x_i < x_j$ implies $y_i - x_i < y_j - x_j$. This property is familiar to a monotonicity property in fair division literature, see Thomson [2003].

[5] We assume that $x >> 0$ when discussing transfer maneuvers among three or more users, in Section 6.

risk neutral, so that each nonnull user wants to minimize her (expected) wait time until completion of her own job. We assume that the system manager can punish those users who exaggerate their demands, so that each user reveals her true job size.

As any deterministic scheduling discipline is vulnerable to either splitting or merging (Moulin (2004b)), we consider the randomization of the ordering of jobs. Note that, unlike the section on coordinated split-proofness, the section on transfer-proofness assumes that, after the server draws the ex post ordering of the jobs, the users acquire this information.

Let $\Phi(N)$ denote the set of orderings of N, with generic element $\sigma : N \to \{1, ..., |N|\}$, interpreted as i is served before j if $\sigma(i) < \sigma(j)$. Given σ, the wait time of user i, denoted by y_i^σ, is

$$y_i^\sigma = x_i + \sum_{j:\sigma(j)<\sigma(i)} x_j$$

In order to extend this definition to random orderings on N, let p be a probability distribution on $\Phi(N)$, and denote by $\Delta(\Phi(N))$ the set of such distributions. Now, given a scheduling problem (N, x) and a random ordering p, the expected wait of a user is given by

$$y_i = \sum_{\sigma\in\Phi(N)} p(\sigma) \cdot y_i^\sigma = x_i + \sum_{j\in N\setminus i} prob\{\sigma(j) < \sigma(i)\} \cdot x_j \tag{1}$$

where $prob\{\sigma(j) < \sigma(i)\} = \sum_{\substack{\sigma\in\Phi(N):\\ \sigma(j)<\sigma(i)}} p_\sigma$.

We say that a wait profile $y \in \mathbb{R}_+^N$ is feasible at a given problem (N, x) if equation (1) holds for some lottery $p \in \Delta(\Phi(N))$. Queyranne (1992) defines the set of feasible wait profiles. Given a scheduling problem (N, x), define for all $S \subset N$, the function $v(S) = \sum_S x_i^2 + \sum_{i,j\in S, i\neq j} x_i x_j$. Note that v is a supermodular function. Now, y is feasible if and only if $i)$ $x_i = 0$ implies $0 \leq y_i \leq \sum_{j\in N} x_j$ and $ii)$ $x \cdot y$ belongs to the core of the game $(N, v(\cdot, x))$, i.e.

$$\sum_{i\in N} x_i \cdot y_i = v(N, x) \text{ and } \sum_{i\in S} x_i \cdot y_i \geq v(S) \text{ for all } S \subset N.$$

This result implies that all random orderings $p \in \Delta(\Phi(N))$ are efficient, since the weighted sum $\sum_{i\in N} x_i \cdot y_i$ is independent of p. See also Moulin (2004b).

Definition 1 *A scheduling rule is a mapping p associating to each problem $(N,x), x \in \mathbb{R}^N_{++}$, a random ordering $p(N,x) \in \Delta(\Phi(N))$. A scheduling method is a mapping y associating to each problem $(N,x), x \in \mathbb{R}^N_{+}$, a feasible wait profile $y(N,x)$.*

If the scheduling problem (N,x) includes some null jobs, we extend the definition of a scheduling rule by saying that it serves the null jobs first. Since all users are indifferent as to when a null user leaves the queue, this simplifies the analysis without any loss of generality.

A key scheduling method is the *Uniform method.* Given (N,x), to each nonnull user i, this method assigns the expected wait $y_i^{Uni}(N,x) = x_i + \frac{1}{2}\sum_{j \in N \setminus i} x_j$. One scheduling rule that implements the Uniform method is the Uniform rule which assigns equal probability to all the orderings of N. There are many scheduling rules that implement the Uniform method. Given (N,x), and $\sigma \in \Phi(N)$, let $\bar{\sigma}$ denote the inverse ordering of σ: for all $i,j \in N$, $\bar{\sigma}(i) < \bar{\sigma}(j)$ iff $\sigma(j) < \sigma(i)$. Given a scheduling rule p, define $\bar{p}$ as the "inverse" of p: for all $\sigma \in \Phi(N)$ we have $\bar{p}(\sigma) = p(\bar{\sigma})$. The scheduling rule $\frac{1}{2}p + \frac{1}{2}\bar{p}$ implements the Uniform method.

Given (N,x), we write $x_{[S]}$ for the restriction of x to $S \subset N$. We also write $x \geq 0$ if $x_i \geq 0$ for all $i \in N$ and $x \gg 0$ if all coordinates of x are positive.

1.4 Normative Properties of Scheduling Rules

We introduce some desirable properties for scheduling rules and methods. In order to economize on space, when a certain property is defined for a scheduling method, we say that a scheduling rule satisfies this property if its associated method does.

A mild fairness property for scheduling rules is anonymity, which requires that the scheduling rule be independent of the naming of the users.

Definition 2 *Given $(N,x), x \gg 0$, and a bijection $\tau : N \to N$ let (N,x^τ) denote the scheduling problem where $\forall i \in N : x_i^\tau = x_{\tau(i)}$. The scheduling rule p satisfies anonymity if , for all $\sigma \in \Phi(N)$*

$$p_{\tau\circ\sigma}(N,x^\tau) = p_\sigma(N,x)$$

Given $(N,x)\,, x \gg 0$, *the scheduling method* y *satisfies anonymity if, for all* $i \in N$

$$y_i\left(N, x^{\tau}\right) = y_{\tau(i)}\left(N, x\right)$$

If a scheduling rule satisfies anonymity, its associated method does so as well. However, it is possible to implement an anonymous method by a nonanonymous rule. For an example of such a rule, see the discussion after Theorem 1.

As a user would have a wait of x_i if nobody else was present, we interpret $y_i - x_i$ as the externality imposed on user i by the other users. When x_i increases, the externality that user i imposes on other users increases as well. A desirable property of a scheduling method is that $y_i - x_i$ should be nondecreasing in x_i.

Definition 3 *The scheduling method* y *satisfies demand monotonicity if for all* $(N,x)\,, x \gg 0$, *and for all* $i \in N$, $y_i\left(N,x\right) - x_i$ *is nondecreasing in* x_i.

Under a demand monotonic scheduling method, as the job size of i decreases, her net wait, $y_i\left(N,x\right) - x_i$, will be weakly decreasing. If x_i is very small compared with the other jobs in x, it might be in our interest to serve x_i very quickly. This concern is the subject of the following property for scheduling methods.

Definition 4 *The scheduling method* y *satisfies small job preference if, for all* (N,x),$x \gg 0$, *for all* $i \in N$, *and for all sequences* a^k *such that* $\lim_{k\to\infty} a^k = 0$, *we have* $\lim_{k\to\infty} y_i\left(N, x^k\right) = 0$, *where* $x_i^k = a^k$ *and* $x_j^k = x_j$ *for all* $j \neq i$.

Separability requires that the relative ordering of a group of users S does not depend on the jobs of the users outside S.

Definition 5 *(Moulin (2004b)) The scheduling rule* p *satisfies separability if for all* $(N,x)\,, x \gg 0$, *and for all* $S \subset N$, *the random ordering of the jobs in* S *is independent of the jobs outside* S :

for all $\sigma^* \in \Phi\left(S\right)$: $\quad \sum_{\sigma\in\Phi(N):\sigma[S]=\sigma^*} p_\sigma\left(N,x\right)$ *is independent of* $x_{[N\setminus S]}$.

Applying the above definition for $S = \{i, j\}$, we see that for a separable rule, the random ordering of the jobs x_i, x_j will be independent from $x_{[N\setminus S]}$. Consequently, given N, we can speak of the probability that a job of size x_j for user j will precede a job of size x_i for user i. We write this probability as $\theta^{i,j,N}(x_i, x_j)$, and the method y associated with p can be expressed as the following

$$y_i = x_i + \sum_{j \in N\setminus i} \theta^{i,j,N}(x_i, x_j) \cdot x_j \tag{2}$$

Definition 6 *(Moulin (2004b)) The scheduling method y satisfies separability if, for all $N \subset \mathcal{N}$, $i, j \in N$, there exists a function $\theta^{i,j,N} : \mathbb{R}^2_+ \to [0,1]$, such that for all $a, b > 0$*

$$\theta^{i,j,N}(a,b) + \theta^{j,i,N}(b,a) = 1, \text{ and } a > 0 \Rightarrow \theta^{i,j,N}(a,0) = 1$$

and such that equation (2) holds for all (N, x).

For a separable method, $\sum_{i\in S} x_i y_i = \sum_{i\in S} x_i \left(x_i + \sum_{j\in N\setminus i} \theta^{i,j,N}(x_i, x_j) \cdot x_j\right) = \sum_S x_i^2 + \sum_{S(2)} x_i x_j = v(S)$ since $\theta^{i,j,N}(a,b) + \theta^{j,i,N}(b,a) = 1$. Hence, $x \cdot y$ is in the core of $(N, v(\cdot, x))$, implying that any separable method can be implemented by a scheduling rule. However, this scheduling rule does not have to be separable. Again, we refer the reader to the discussion after Theorem 1 for an example of such a rule.

The following subclass of scheduling rules stand out in our analysis.

Definition 7 *(Moulin (2004b)) Let w be a function on $\mathbb{R}_{++}$ such that $w(a) > 0$ for $a > 0$. Given a scheduling problem (N, x) with $x \gg 0$, the w-quasi-proportional rule draws $|N|$ independent random variables Z_i with respective c.d.f.s $F(z_i) = \min\left\{z_i^{w(x_i)}, 1\right\}$ for all $z_i \geq 0$, and schedules jobs in the order of the realized Z_i's, breaking ties randomly with uniform probability.*

For a characterization of $w-$quasi-proportional rules, see Moulin (2004b). As he notes, a $w-$quasi-proportional scheduling rule satisfies anonymity as the c.d.f. is only a function of job size, and not the names of the users; it satisfies separability since the random variables Z_i are

independent. The associated method of a $w-$quasi-proportional rule takes the form

$$y_i(N,x) = x_i + \sum_{j \in N \setminus i} \theta(x_i, x_j) \cdot x_j$$

where

$$\theta(x_i, x_j) = \frac{w(x_i)}{w(x_i) + w(x_j)}$$

Our next axiom is scale invariance, which requires that a scheduling rule be independent of the units we measure the jobs in.

Definition 8 *The scheduling method y satisfies scale invariance if, for all (N,x) with $x \geq 0$ and for all $\gamma > 0$, we have $y(N, \gamma \cdot x) = \gamma \cdot y(N,x)$. The scheduling rule ρ satisfies scale invariance if the associated method does.*

Lemma 1 *A w-quasi-proportional rule satisfies scale invariance and demand monotonicity if and only if $w(x) = Cx^{\alpha}$ for some $C, \alpha \geq 0$.*

Proof. See the appendix.

Note that, the $w-$quasi-proportional rule with $w(x) = Cx^{\alpha}$ is the same as the one with $w(x) = x^{\alpha}$. Accordingly, we restrict attention to the class of $w-$quasi-proportional rules with $w(x) = x^{\alpha}$, and call this family of scheduling rules the *power family.* Let p^{α} be a generic element of this family. When $\alpha = 0$, we get the Uniform rule. As we increase α, we see that the probability of serving a job of size x before a slightly longer job increases. We get the Shortest Job First rule as a limiting case of quasi-proportional scheduling rules as α tends to ∞. Likewise, as α tends to $-\infty$, we get the opposite *Longest Job First (LJF) rule.*

An important quasi-proportional rule is the *Proportional rule*, for which $\alpha = 1$. Under this rule, total externality is divided in a proportional manner between two users: $\theta(a,b) = \frac{a}{a+b}$. The Proportional rule stands out among split-proof rules in that it performs well in terms of the tests we outlined in the introduction. However, as we show in the next section, it is vulnerable to coordinated splitting.

1.5 Coordinated Split-Proofness

The following example illustrates the idea of coordinated splitting among two users.

Example: Pairwise Splitting Under the Proportional Rule

To each nunnull user i, the Proportional method assigns the expected wait

$$y_i(N,x) = x_i + \sum_{j \in N \setminus i} \frac{x_i}{x_i + x_j} \cdot x_j$$

Consider the problem $N = \{A, B, C\}$ and $x = (2, 2, 1)$. We first compute the expected delays of A and B, denoted by $y_A = y_B$, under the Proportional method:

$$\begin{aligned} y_A = y_B &= x_A + \theta^{AB}(x_A, x_B) \cdot 2 + \theta^{AC}(x_A, x_C) \cdot 1 \\ &= 2 + \tfrac{1}{2} \cdot 2 + \tfrac{2}{3} \cdot 1 = \tfrac{11}{3} \end{aligned}$$

Now, set $S = \{A, B\}$. Suppose that users A and B assume two aliases each, $\{A1, A2\}$ and $\{B1, B2\}$ respectively, and report a job of size 1 under each of these aliases. We write $N_* = \{A1, A2, B1, B2, C\}$ and $x_* = (1, 1, 1, 1, 1)$. Consider the agreement between A and B under which A precedes B in all the orderings, i.e. A uses the first two 1-unit slots that belong to S. For instance, if the ex post ordering of the jobs in x_* turns out to be $B2, C, A1, A2, B1$, user A will use the first two slots $B2$ and $A1$, and user B will use the slots $A2$ and $B1$. Denote by y_A^{AB} and y_B^{AB} the corresponding wait times of A and B. Note that in the new problem, as the Proportional rule is anonymous and all the jobs are of equal size, all of the orderings are equiprobable.

$y_A^{AB} = x_A + prob\{C \text{ is served first or second}\} \cdot x_C = 2 + \frac{2}{5} \cdot 1 = \frac{12}{5}$

$y_B^{AB} = x_B + x_A + prob\{C \text{ is not served last}\} \cdot x_C = 2 + 2 + \frac{4}{5} \cdot 1 = \frac{24}{5}$

Note that, as $\frac{1}{2}\left(\frac{12}{5} + \frac{24}{5}\right) < \frac{11}{3}$, the agreement under which A precedes B with probability $\frac{1}{2}$ is profitable for both of these users.

We do not have this problem under the Uniform rule. Under this rule, the wait time of user C in the original problem is $y_C^{Uni} = 1 + 2 \cdot \left(\frac{1}{2} \cdot 2\right) = 3$. In the new problem, her wait time is still

$1 + 4 \cdot \left(\frac{1}{2} \cdot 1\right) = 3$. As the Uniform rule is efficient, this means users A and B cannot both benefit from the above coordinated splitting maneuver.

The following notation will be useful in defining coordinated split-proofness. Given a scheduling problem (N, x), take $S \subset N$. Suppose each user i in S splits into $i \cup T_i$, $T_i \cap N = \varnothing$, and $T_i \cap T_j = \varnothing$ for all $i, j \in S, i \neq j$. Set $T = \cup_{i \in S} T_i$. The new problem becomes (N^*, x^*) with $N^* = N \cup T$, $\sum_{k \in i \cup T_i} x_k^* = x_i$ for all $i \in S$ and $x_j^* = x_j$ for all $j \in N \backslash S$.

We now define the set of feasible wait times achievable by S under their reported demands. The users in S are free to decide on a possibly randomized ordering according to which they will serve their "true" jobs using the split slots $(x^*)_{[T]}$.

Definition 9 *We say that a group of users S agree on $\tau \in \Phi(S)$ if for all $i, j \in S$, user i precedes user j iff $\tau(i) < \tau(j)$. We say that they agree on $q \in \Delta(\Phi(S))$ if they implement agreement τ with probability q_τ for all $\tau \in \Phi(S)$.*

In other words, the users in S agree on τ if, the first user according to τ uses the first slots that are available to S until the completion of her job, the second user according to τ uses the next slots available to S until the completion of her job, etc.

Given an agreement q among S, we now compute the expected wait of a user $i \in S$ induced by this agreement, denoted by y_i^q. Let $P(i; \sigma, \tau)$ denote the set of users in N that precede i under σ and τ : $P(i; \sigma, \tau) = \{j \in S : \tau(j) < \tau(i)\} \cup \{j \in N \backslash S : \sigma(j) < \sigma(i_* | i; \sigma, \tau)\}$, where $\{i_* | i; \sigma, \tau\} = \arg\min_{j \in T} \sigma(j)$ such that $\sum_{\substack{k \in T: \\ \sigma(k) \leq \sigma(j)}} (x^*)_k \geq \sum_{\substack{k \in S: \\ \tau(k) \leq \tau(i)}} x_k$. In other words, i_* is the alias under which user i completes her job. We have

$$y_i^q = x_i + \sum_{\tau \in \Phi(S)} q_\tau \cdot \sum_{\sigma \in \Phi(N_*)} p_\sigma \cdot \sum_{j \in P(i; \sigma, \tau)} x_j$$

Definition 10 *Take a scheduling rule p and a scheduling problem (N, x), $S \subset N$. We say that p is vulnerable to coordinated splitting by S if, there exist $\{T_i\}_{i \in S}$ and x_* defined as above, and $q \in \Delta(\Phi(S))$ such that*

$$\forall i \in S : \; y_i^q < y_i$$

We say that p is coordinated split-proof (CSP) if, for all scheduling problems (N,x) and for all $S \subset N$, p is invulnerable to coordinated splitting by S.

We say that p is pairwise split-proof (PSP) if, for all scheduling problems (N,x) and for all $S \subset N$ such that $|S| = 2$, p is invulnerable to coordinated splitting by S.

If in the above definition we restrict $|S| = 1$, we get the definition of split-proofness as it appears in Moulin (2004b). There is a wide class of scheduling rules that are split-proof. The Proportional rule is among these rules. However, as we saw in the example above, it is not pairwise split-proof. The following lemma shows that many of the scheduling rules that are split-proof are vulnerable to pairwise splitting.

Lemma 2 *Among anonymous scheduling rules, there is no rule that satisfies small job preference and pairwise split-proofness.*

Proof. Take an anonymous rule p that satisfies small job preference, and let y be the scheduling method associated with p. We show that p is vulnerable to pairwise splitting. To this purpose, for each integer $k \geq 2$, define the problem $\left(N, x^k\right)$ with $N = \{1,2,3\}$ and $x = \left(1, 1, \frac{1}{k}\right)$. Let $y^k = y\left(N, x^k\right)$. We have

$$y_1^k = y_2^k = 1 + \frac{1}{2} + \text{Prob}\left\{\frac{1}{k} \text{ is served before } 1\right\}\frac{1}{k}$$

For each $k \geq 2$, consider a pairwise splitting maneuver by users 1 and 2, where each of the users splits into k users and reports a demand of $\frac{1}{k}$ under each of these aliases. Consider an agreement between users 1 and 2 where user 1 is served before user 2 with probability $\frac{1}{2}$. For all $k \geq 2$, denote by $\left(y^k\right)^*$ the expected wait induced by this agreement. Note that after the splitting move, we have $2k+1$ jobs of size $\frac{1}{k}$. As p is anonymous, all the orderings are equally likely in the resulting problem. Hence, the probability that user 3 is served before user 1 (or user 2) under the specified agreement is $\frac{1}{2}\frac{k}{2k+1} + \frac{1}{2}\frac{2k}{2k+1} = \frac{3k}{4k+2}$. This implies that

$$\left(y^k\right)_1^* = \left(y^k\right)_2^* = 1 + \frac{1}{2} + \frac{3k}{4k+2}\frac{1}{k}$$

Now, since p satisfies small job preference, there exists K big enough such that Prob$\{\frac{1}{K}$ is served before $1\} > \frac{3K}{4K+2}$ and hence the above agreement is beneficial for users 1 and 2 in the problem (N, x^K). ■

The above lemma shows that pairwise split-proofness is a very strong property. To find CSP scheduling rules we relax the small job preference property. Any scheduling rule that implements the Uniform method, for instance, is CSP. To see this, take any (N, x) and fix $S \subset N$. Recall that under the Uniform method, the expected delay of a user is $y_i^{Uni} = x_i + \frac{1}{2}\sum_{j \in N \setminus i} x_j$. Hence, for a user i outside S, the sum $\sum_{j \in N \setminus i} x_j$ is unaffected by a coordinated splitting move by S. As our scheduling rule is efficient, this implies the users in S cannot all benefit from a coordinated splitting maneuver.[6]

The Longest Job First rule is also CSP by a similar argument. Under this rule, a user outside S either is indifferent to a coordinated splitting maneuver by S or benefits from it. In a similar fashion, for $\alpha \in [0, 1]$, one checks that all the rules of the form $\alpha UNI + (1 - \alpha) LJF$ are CSP. Note that for $\alpha < 1$, none of these rules are demand monotonic. The next result characterizes scheduling rules that are demand monotonic and CSP among separable and anonymous scheduling rules.

Theorem 1 *Assume that $|\mathcal{N}| = \infty$. In the class of separable and anonymous scheduling rules, the rule p is demand monotonic and CSP if and only if its corresponding method is Uniform.*

Proof. We already showed above that any scheduling rule that implements the Uniform method is CSP. One checks that it is demand monotonic as well.

To show the only if part, let p be a separable scheduling rule that satisfies the premises of the above theorem, and let y be the scheduling method associated with p. Since p satisfies *separability*, y is of the form (2). As p is *anonymous*, so is y. So, given $N \subset \mathcal{N}$, and $i, j \in N$, we have $\theta^{i,j,N} = \theta^N$.

We first outline the steps that we will follow in the proof of the theorem: the initial step is to show that θ must be independent of N. The next step is to show that $\theta(a, b) = 1/2$ for all $a, b > 0$ in order to establish that y is the Uniform method. *Demand monotonicity* implies that $\theta(a, b)$ is

[6] One could formulate a stronger CSP property by changing the condition in Definition 10 to read $\forall i \in S : y_i^q \leq y_i$ and $\exists j \in S : y_j^q < y_j$.
The Uniform rule satisfies this alternative form of coordinated split-proofness.

nondecreasing in a for all $b > 0$. Hence, for any $z > 0$, and all integers $K \geq 2$, *demand monotonicity* implies $\theta(Kz, z) \geq 1/2$. To show that $\theta(Kz, z) \leq 1/2$, we follow a method that closely follows the spirit of Example 1: there are several users with jobs of equal size who will be making a coordinated splitting move. Suppose that there is some agreement between these users that is beneficial for them. We observe that the symmetric agreement which orders all the jobs in the manipulating coalition with equal probability must be a beneficial agreement as well. So, we direct our attention to the symmetric agreement and show that, if there are enough users in the manipulating coalition, *coordinated split-proofness* implies $\theta(Kz, z) \leq 1/2$. This establishes that $\theta(Kz, z) = 1/2$ for all $z > 0$ and $K \geq 2$. Since $\theta(a, b)$ is nondecreasing in a for all $b > 0$, we must have $\theta(a, b) = 1/2$.

We first show that θ is independent of N, i.e. for all $N, \bar{N} \in N$, $\theta^N(a, b) = \theta^{\bar{N}}(a, b)$. Pick $N_1, N_2 \in \mathcal{N}$ such that $S = \{i_1, i_2\} \subset N_1 \subset N_2$. Given any $a, b > 0$, suppose $\theta^{N_1}(a, b) \neq \theta^{N_2}(a, b)$. Since $\theta^N(a, b) + \theta^N(b, a) = 1$ for all $N \in \mathcal{N}$, we can, without loss of generality, assume that $\theta^{N_1}(a, b) > \theta^{N_2}(a, b)$. Consider the problem (N_1, x), where $x_S = (a, b)$ and $x_i = 0$ for all $i \neq i_1, i_2$. User i_1 can benefit by splitting into $i_1 \cup (N_2 \backslash N_1)$, reporting a job of size zero under all his aliases $N_2 \backslash N_1$. This contradiction establishes that $\theta^{N_1}(a, b) = \theta^{N_2}(a, b)$. Hence, for all $N, \bar{N} \in \mathcal{N}$ we have $\theta^N(a, b) = \theta^{N \cup \bar{N}}(a, b) = \theta^{\bar{N}}(a, b)$, and θ is independent of N.

We now construct an example that shows y must be the Uniform method, i.e. $\theta(a, b) = 1/2$ for all $a, b > 0$. To this purpose, we generalize the idea in Example 1. Take any $z > 0$ and any integer $K \geq 2$. Consider the problem (N, x) with $N = \{1, 2, ..., n+1\}$, $x_{n+1} = z$ and $x_i = Kz$ for all $i = 1, 2, ..., n$. Set $S = \{1, 2, ..., n\}$ and denote by $\tilde{y} = (\tilde{y}_i)_{i \in N}$ the expected wait profile in the problem (N, x).

$$\begin{aligned} \forall i \in S: \quad \tilde{y}_i &= Kz + (n-1)\theta(Kz, Kz)Kz + \theta(Kz, z)z \\ &= \tfrac{(n+1)}{2}Kz + \theta(Kz, z)z \end{aligned}$$

Note that, since p is *demand monotonic*, so is y. Take any $a, b > 0$. In the problem $(\{1, 2\}, (a, b))$ demand monotonicity implies $y_1 - a = \theta(a, b)b$ is nondecreasing in a, implying that $\theta(a, b)$ is nondecreasing in a. Since $\theta(z, z) = 1/2$ by *anonymity* and $K \geq 2$, this in turn implies $\theta(Kz, z) \geq 1/2$.

We now show $\theta(Kz, z) \leq 1/2$. Suppose each of the users in S splits her demand into K equal pieces, so that in the new problem there are $Kn+1$ users, each with a demand of size z. Note that, by *anonymity* of the rule p, all orderings are equiprobable in the problem after splitting. Next, consider the agreement $\bar{\tau}$ among S, where $\bar{\tau}(i) = i$ for all $i \in S$. Denote by $y^{\bar{\tau}} = (y_i^{\bar{\tau}})_{i \in S}$ the true expected delays of the users in S induced by the agreement $\bar{\tau}$. We have

$$\begin{aligned} y_i^{\bar{\tau}} &= iKz + prob\{\text{user } (n+1) \text{ is served at most in } Ki \text{ th place}\} z \\ &= iKz + \tfrac{iK}{Kn+1} z \end{aligned}$$

We now consider the agreement that randomizes on τ, and denote by y^* the corresponding wait time for a user in S

$$\begin{aligned} y^* &= \tfrac{1}{n} \textstyle\sum_{i=1}^{n} y_i^{\bar{\tau}} = \tfrac{1}{n} \sum_{i=1}^{n} \left(iKz + \tfrac{iK}{Kn+1} z \right) \\ &= \tfrac{(n+1)}{2} Kz + \tfrac{Kn+K}{2Kn+2} z \end{aligned}$$

For p to be *coordinated split-proof*, we must have $\tilde{y}_i \leq y^*$ for all $i \in S$, which implies

$$\tfrac{(n+1)}{2} Kz + \theta(Kz, z) z \leq \tfrac{(n+1)}{2} Kz + \tfrac{1}{2} z$$

Therefore, $\theta(Kz, z) \leq 1/2$ for all $z > 0$ and for all integers $K \geq 2$.

We showed that $\theta(Kz, z) = 1/2$ for all $z > 0$ and for all integers $K \geq 2$. Since $\theta(a, b)$ is nondecreasing in a, this implies that $\theta(a, b) = 1/2$ for all $a, b > 0$. Hence, y is the Uniform method.

■

Within the class of separable and anonymous scheduling rules, demand monotonicity and coordinated split-proofness are logically independent properties. Recall that the Longest Job First rule is CSP. It is also anonymous and separable. However, it is not demand monotonic as increasing one's job size might put her ahead in the queue. The proportional rule is an example of a separable and anonymous scheduling rule that is not CSP but is demand monotonic.

In the proof of Theorem 1, we use the assumption that $x \in \mathbb{R}_+^N$ in order to show that θ is independent of N. What would happen if we restricted our problems so that $x \in \mathbb{R}_{++}^N$? It turns out that a *population monotonicity* property, in the spirit of Thomson (1983), is enough to guarantee

us the independence of θ from N.

Definition 11 *The scheduling method y is population monotonic if, for all (N,x), for all $i \in N$, $j \notin N$, and for all $a \geq 0$, $y_i(N,x) \leq y_i\left(N \cup j, x^j\right)$, where $x_i^j = x_i$ for all $i \in N$ and $x_j^j = a$.*

Restricting ourselves to the domain of strictly positive job profiles, we have the following result.

Theorem 1 (Alternative): The anonymous, separable, and population monotonic scheduling rule p satisfies demand monotonicity and CSP if and only if its associated method is Uniform.

Proof. Let p be a separable scheduling rule that satisfies the premises of the above theorem, and let y be the scheduling method associated with p. Since p satisfies *separability*, y is of the form (2). As p is *anonymous*, so is y. So, given $N \subset \mathcal{N}$, and $i,j \in N$, we have $\theta^{i,j,N} = \theta^N$. Thanks to our argument in the proof of Theorem 1, it suffices to show that θ is independent of N.

Take any $i_1, i_2 \in \mathcal{N}$ and any $N_1, N_2 \subset \mathcal{N}$ such that $\{i_1,i_2\} \subset N_1 \subset N_2$. Suppose that $\theta^{N_1}(a,b) \neq \theta^{N_2}(a,b)$ for some $a,b > 0$. Since $\theta^N(a,b) + \theta^N(b,a) = 1$ for all $N \in \mathcal{N}$, we can, without loss of generality, assume that $\theta^{N_1}(a,b) > \theta^{N_2}(a,b)$. Consider the problem (N_1,x) where $x_{i_1} = a, x_{i_2} = b$ and $x_{i_k} = c$ for all $i_k \in N_1 \backslash \{i_1,i_2\}$, and the problem $(N_1,\bar{x})$, where $\bar{x}_{i_1} = a, \bar{x}_{i_2} = b$ and $\bar{x}_{i_k} = c$ for all $i_k \in N_2 \backslash \{i_1,i_2\}$. If c is sufficiently close to 0, we have $y_1(N_2,\bar{x}) = a + \theta^{N_2}(a,b)\,b + (|N_2| - 2)\,\theta^{N_2}(a,c)\,c < a + \theta^{N_1}(a,b)\,b + (|N_1| - 2)\,\theta^{N_1}(a,c)\,c = y_1(N_1,x)$, which is a violation of *population monotonicity*. This implies $\theta^{N_1}(a,b) = \theta^{N_2}(a,b)$ for all $a,b > 0$.

Hence, for all $N, \bar{N} \subset \mathcal{N}$ we have $\theta^N(a,b) = \theta^{N \cup \bar{N}}(a,b) = \theta^{\bar{N}}(a,b)$, and θ is independent of N.

■

The problem with this theorem is that, population monotonicity might be a property that is implied by CSP in the presence of the other properties listed in the theorem. The difficulty of finding CSP scheduling rules that are dependent on N makes this a challenging task.

1.6 Transfer-Proofness

We now turn to the transferring of jobs between users. Given a scheduling problem (N,x), and $S \subset N$, suppose the users in S transfer their jobs among themselves. Then, the new problem

becomes (N, x_*), where $\sum_{i\in S}(x_*)_i = \sum_{i\in S} x_i$ and $(x_*)_j = x_j$ for all $j \in N\backslash S$. In other words, the users in S report a new allocation of their total demand $\sum_{i\in S} x_i$.

The server will now assign S the slots $(x_*)_{[S]}$.The users in S are free to choose a possibly randomized ordering according to which they would serve their "true" jobs in these slots. Recall that in the previous section, we interpreted an agreement among a group of users S as a (random) priority ordering on the usage of the slots available to S. We now assume that the users can make a more sophisticated agreement, under which they can condition this priority ordering on the ex post ordering of the jobs in x_*.

Consequently, an agreement among S is a function q associating to each ordering σ in $\Phi(N)$ a random ordering $q(\sigma) \in \Delta(\Phi(S))$. Given a scheduling rule p and an agreement q among S, the true wait time of a user $i \in S$ induced by this agreement, denoted by y_i^q, is given by

$$y_i^q = x_i + \sum_{\sigma\in\Phi(N)} p_\sigma \cdot \sum_{\tau\in\Phi(S)} q_\tau(\sigma) \cdot \sum_{j\in P(i;\sigma,\tau)} x_j$$

where we define $P(i,\sigma,\tau)$ as in Section 5, with $T_i = \varnothing$ for all $i \in S$.

Definition 12 *Take a scheduling rule p, a scheduling problem (N,x), and $S \subset N$. We say that p is vulnerable to transferring by S if, there exists x_* defined as above and $q : \Phi(N) \to \Delta(\Phi(S))$ such that*

$$\forall i \in S:\ y_i^q \leq y_i \text{ and } \exists j \in S:\ y_j^q < y_j$$

We say that p is transfer-proof if, for all scheduling problems (N,x), and for all $S \subset N$, p is invulnerable to transferring by S.

Definition 13 *We say that the scheduling rule p is pairwise transfer-proof if, for all scheduling problems (N,x) and for all $S \subset N, |S| = 2$, p is invulnerable to transferring by S.*

Transfer-proofness is very close to the axiom of *no advantageous reallocation* considered in the literature. Moulin (1985) proposed it in the problem of dividing social surplus and Chun (1988) uses it in the context of *rights problems*. In both of these papers, the proportional rule satisfies the

no advantageous reallocation property. The theorem below gives a class of scheduling rules that are pairwise transfer-proof, among them the proportional rule.

Theorem 2 *A w–quasi-proportional scheduling rule p^w is pairwise transfer-proof if w is concave and nondecreasing.*

We leave the proof to the appendix.

Recall that the power family is the set of $w-$quasi-proportional scheduling rules for which $w(x) = x^\alpha$ for some $\alpha \geq 0$. The above theorem suggests that a scheduling rule from the power family is pairwise transfer-proof if $\alpha \in [0,1]$. The third part to the proof of Proposition 1 below shows that p^α is not transfer-proof for any $\alpha > 1$.

We now allow transfers between three or more users. Note that, in the presence of null jobs, transferring maneuvers would effectively grant the possibility of splitting to a user. Hence, a negative result similar to Theorem 1 would apply to transfer-proofness. To see if another impossibility result can be avoided, we restrict the domain of scheduling problems so that the job profile is strictly positive. This is a reasonable assumption in many queueing situations, where users arrive at the server only if they have a positive job to be processed. Indeed, whenever the users have to incur a fixed cost to arrive at the server (e.g. take the time to go to the bank), this assumption will be satisfied. Our next result suggests that, within the power family, we cannot avoid an impossibility result.

Proposition 1 *Assume that $|\mathcal{N}| \geq 4$. A scheduling rule from the power family is transfer-proof if and only if its associated method is Uniform.*

We leave the proof to the appendix.

Even though the proof of this result uses the "strong" version of transfer-proofness that we introduced in Definition 12, the results remain valid if we employ a weaker version of transfer-proofness, replacing the weak vector dominance in this definition with strong vector dominance.

A similar result to Proposition 1 appears in the context where monetary transfers are feasible: there is a large class of mechanisms immune to transfers involving two users, but any "continuous"

mechanism, i.e. a mechanism under which total waiting time is a continuous function of the job profile x, is vulnerable to transfers involving three or more users (Moulin (2004a)).

1.7 Concluding Remarks

We developed an argument in favor of the Uniform method based on its incentive properties. However, several criticisms can be directed towards this method. One of the immediate criticisms of the Uniform method is that the sum of the expected waits of all the users, the total waiting time, can be infinitely longer than necessary. The Uniform method does not meet ranking: $x_i < x_j$ implies $y_i^{Uni} - x_i > y_j^{Uni} - x_j$. Also, under this method, if all the other users are big, a very small job might have a very big expected wait.

In light of these criticisms, our results on coordinated split-proofness and transfer-proofness may be seen as negative results. We know that the Longest Job First rule also satisfies CSP, but among others, all the criticisms of the Uniform rule also apply to this rule. Perhaps if we relaxed separability we would be able to find scheduling rules that satisfy CSP and have "better" properties than the Uniform rule. Likewise, other transfer-proof rules might be found outside the class of quasi-proportional rules. The complicated nature of coordinated splitting and transferring maneuvers makes this a challenging question.

2 A Characterization of Bird's Rule

We consider covert-merging and overt-merging as forms of strategic behavior in minimum cost spanning tree games. There is no covert-merge–proof cost allocation rule. Several rules studied in the literature are vulnerable to overt-merging as well. *Bird's rule*, on the other hand, is characterized by overt-merge–proofness, core selection, and tree invariance.

2.1 Introduction

We study strategic behavior in the form of merging in minimum cost spanning tree (mcst) games: a good or service is to be distributed from a source to a fixed set of users, and we are interested in allocating the cost of the distribution network. Mcst games arise in various economic situations like the construction of electricity, cable, and irrigation networks. Since their introduction into economics literature by Claus and Kleitman (1973), several papers have studied the problem, focusing on different issues like core stability and the structure of the core (Granot and Huberman, 1981, 1982, 1984; Feltkamp et. al., 1994), fairness properties like equal treatment, cost monotonicity and population monotonicity (Norde et.al., 2001; Kar, 2002; Bergantiños and Vidal-Puga, 2004), and consistency (Feltkamp et. al., 2000; Dutta and Kar, 2004).[7] Claus and Kleitman (1973) raised the issue of finding cost allocation rules that are immune to merging maneuvers by the users.

We consider two different types of merging: *covert-merging* and *overt-merging*. We use an example to illustrate the difference between these two concepts.

2.1.1 An illustrative example

Consider the mcst problem P presented in Figure 1. $N = \{A, B, C\}$ is the set of users and 0 is a special node called *the source*, e.g. a power plant or a dam. The numbers alongside the edges in the graph are the costs of establishing those edges. The cheapest network spanning all the users, i.e. the mcst, is indicated with a heavy line.

[7] See Sharkey (1995) for a survey of network models in economics.

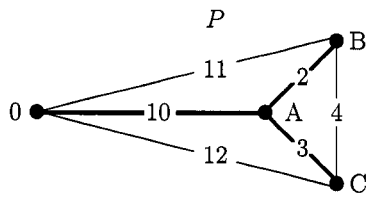


Figure 1

A *covert-merging maneuver by coalition* $S = \{A, C\}$ *into* C results in the mcst problem P^C of Figure 2a. User A disappears from the problem, and obtains access to the source *only* through user C via a private network, the costs of which are borne by coalition S. Note that the private network is hidden. From the point of view of the service provider, it is as if user A does not exist. We assume that S will build the cheapest private network, i.e. a mcst spanning the members of S. In this case, the mcst on S is comprised of the link (AC), which costs 3.

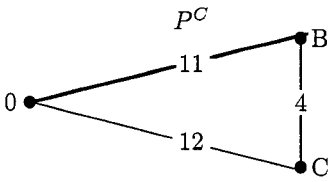


Figure 2a
$\{A, C\}$ covert-merges into C

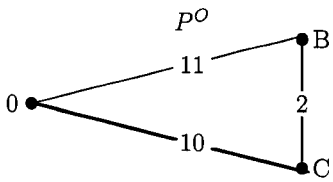


Figure 2b
$\{A, C\}$ overt-merges into C

An *overt-merging maneuver* by S, on the other hand, gives rise to the mcst problem P^O of Figure 2b. In this case, the service provider is aware of the fact that user C is connected to user A through the private network. Consequently, when it wishes to connect the merged user C to the source, it can use the cheaper one of the links $(0A)$ and $(0C)$. This implies that the cost of the link $(0C)$ in problem P^O is 10.

For each $i \in N$, let $\alpha(i)$ denote the immediate predecessor of i in the mcst, e.g. $\alpha(C) = A$ in P and $\alpha(C) = 0$ in P^O. *Bird's rule* (Bird, 1976) charges each user the cost of the link $(\alpha(i)\ i)$. Bird allocation of P is given by $(10, 2, 3)$. Under Bird's rule, the merged user C pays 4 in P^C and 10 in P^O. Thus, the total cost to $\{A, C\}$ after the covert-merging maneuver is $4 + 3 = 7$, which is less than the original total cost of 13 allocated to this coalition. Hence, Bird's rule is vulnerable to covert-merging. The total cost to $\{A, C\}$ after the overt-merging maneuver, however, stays at 13,

and there is no incentive for S to overt-merge into C.

Another approach to allocating costs is to construct a cooperative game from P and use some well known allocation rules. Following Bird (1976) and most of the literature, we define the worth of a coalition T to be the cost of a mcst on $T \cup 0$, where we can only use the links between the members of $T \cup 0$.[8] This is justified by the assumption that any coalition has free access to the technology to build a network. For example, the worth of coalition $\{B, C\}$ is $11 + 4$, since the mcst on $\{0, B, C\}$ consists of the links $(0B)$ and (BC).

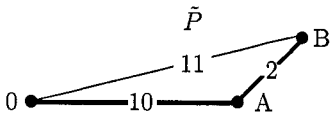


Figure 3a
$\{A, C\}$ covert- or overt-merges into A

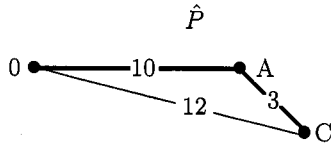


Figure 3b
$\{A, B\}$ covert- or overt-merges into A

Three examples of solutions that are derived from the above cooperative game are the *Shapley value*, the *nucleolus*, and the *Dutta-Kar rule* (Dutta and Kar, 2004).[9] Our results imply that all of these rules are vulnerable to both covert- and overt-merging maneuvers. In the example, the Shapley value charges the cost allocations $Sh(P) = (3.66, 5.17, 6.17)$ .Consider the situation where $S = \{A, C\}$ now merges into A. The problem $\tilde{P}$ of Figure 3a results, regardless of whether the merging is done covertly or overtly. We compute $Sh\left(\tilde{P}\right) = (5.5, 6.5)$. Hence, the $\{A, C\}$ coalition pays $5.5 + 3$ after any merging maneuver, which is less than the original total cost of $3.66 + 5.17$ allocated to this coalition. Similar computations apply to the nucleolus and the Dutta-Kar rule.[10]

[8] The cost sharing game defined in this way may not be monotonic. Megiddo (1978), suggested a monotonic extension of this game, by allowing a coalition T to use the links outside $T \cup 0$. The core of the resulting game is a subset of the core of the original game, and the vulnerability to merging maneuvers persists when we consider solutions derived from such monotone games. See also Granot and Maschler (1998).

[9] The recursive definition of the Dutta-Kar rule is given on page 31.

[10] The *Dutta-Kar rule* charges the cost allocations $DK(P) = (2, 3, 10)$ and $DK\left(\tilde{P}\right) = (2, 10)$. Hence, $S = \{A, C\}$ can decrease their total cost from $2 + 10$ to $2 + 3$ via a (covert- or overt-) merging maneuver.

The *nucleolus* charges $\left(3\frac{1}{3}, 5\frac{1}{3}, 6\frac{1}{3}\right)$ in P. When $\{A, B\}$ covert- or overt-merges into A, problem $\hat{P}$ of Figure 3b arises. The merged user A pays 5.5 in $\hat{P}$, thus decreasing the total cost to $\{A, B\}$ from $8\frac{2}{3}$ to 7.5.

2.1.2 Overview of results

We prove four results. The first one is that there is no *covert-merge–proof* way of allocating costs (Proposition 1). However, if the problem has a unique mcst, Bird's rule is immune to overt-merging maneuvers (Proposition 3). If we allow for multiple mcsts, we are back to an impossibility result regarding *overt-merge–proofness* (Proposition 2). On a restricted domain of cost matrices which ensures a unique mcst, we provide an axiomatic characterization of Bird's rule using *overt-merge–proofness* together with the *core selection* and *tree invariance* properties (see Theorem 1). Core selection requires that a cost allocation rule choose an allocation in the core of the mcst game we defined above, a standard requirement in the mcst literature. Tree invariance, on the other hand, insists that two problems that have the same mcst have the same cost allocation. This strong property is partially justified by the computational ease associated with the rules that satisfy it.

In a contemporaneous paper, Gómez-Rúa and Vidal-Puga (2006) also discuss overt-merge–proofness in a context where the connection costs between towns may be zero. On a similar domain of cost matrices to the one used in our main result, they show that Bird's rule is characterized by overt-merge–proofness, core selection, and *independence of extreme null points*. An extreme null point is a user that has no followers and connects to the network with an edge that has zero cost. Independence of extreme null points asserts that such a user should be charged zero cost and moreover, if we eliminate this user from the problem, the cost allocations of the remaining users should not change. This axiom incorporates an undesirable property of Bird's rule: Suppose the problem looks like that in Figure 3a, except that the cost of the edge (AB) is zero. Then under any cost allocation rule that satisfies independence of extreme null points (Bird's rule in particular), B pays zero and A pays 10, even though B is also using the costly link $(0A)$.

Section 2 introduces the notation and an algorithm due to Jarnik (1930), Prim (1957), and Dijkstra (1959) that constructs a mcst. In section 3 we formalize the two merge-proofness concepts and present the results. Section 4 discusses the implications of the main result and establishes the independence of the axioms.

2.2 Preliminaries

Let $\mathcal{N}$ denote the finite set of potential users. $0 \notin \mathcal{N}$ has the special role of *source* or *root,* who we sometimes refer to as a user for convenience. Given $N \subset \mathcal{N}$, we let $N_0 = N \cup \{0\}$. A *network* on $S \subset \mathcal{N} \cup \{0\}$ is a graph with node set S. A link in this graph is of the form (ij) for some $i, j \in S, i \neq j$. We denote a generic network on S by g_S and the set of connected networks on S by Γ_S .

For a set of users $N \subset \mathcal{N}$, a cost matrix $c = [c_{ij}]_{i,j \in N_0}$ represents the cost of direct connection between users i and j. We assume that $c_{ij} > 0$ if $i \neq j$ and $c_{ii} = 0$ for all $i \in N_0$. The set of all cost matrices for N_0 is denoted by C_N. We let $\mathcal{C} = \cup_{N \subset \mathcal{N}} C_N$.

We assume that all users have to be connected to the network, so we want to construct a connected graph on N_0. Of particular importance are efficient networks, i.e. connected graphs with minimal cost. Such a graph cannot have any cycles, otherwise we could remove an edge from this cycle and still have a (cheaper) connected graph. Hence, the efficient network must be a tree. A minimum cost spanning tree (mcst from now on) over N_0, denoted by $g[N_0, c]$, satisfies $g[N_0, c] = \arg\min_{g \in \Gamma_{N_0}} \sum_{(ij) \in g} c_{ij}$. Also, given $S \subset N_0$, $g[S, c_S]$ denotes the mcst on users S, where c_S is the restriction of c to S. For all $S \subset N_0$, we define $m(S, c_S) = \min_{g \in \Gamma_S} \sum_{(ij) \in g} c_{ij} = \sum_{(ij) \in g[N, c_S]} c_{ij}$.

Given $N \subset \mathcal{N}$ and $c \in C_N$, we call (N, c) a mcst problem. Note that, depending on c, there might be several mcst associated with (N, c). We use two restricted domains. For all $N \subset \mathcal{N}$, let $C_N^1 = \{c \in C_N |$ c induces a unique mcst on $N_0\}$ and $C_N^2 = \{c \in C_N |$all c_{ij} are distinct$\}$. Note that $C_N^2 \subset C_N^1$. We write $\mathcal{C}^i = \cup_{N \subset \mathcal{N}} C_N^i$, for $i = 1, 2$.

Assuming that we constructed an efficient network, we are interested in sharing its cost among the users. This is achieved by a cost allocation rule.

Definition 14 *A cost allocation rule is a family of functions* $\{\phi(N, \cdot)\}_{N \subset \mathcal{N}}$ *such that, for all* $N \subset \mathcal{N}$, $\phi(N, \cdot) : C_N \longrightarrow \mathbb{R}_+^N$ *satisfies* $\sum_{i \in N} \phi_i(N, c) = m(N_0, c)$ *for all* $c \in C_N$.

2.2.1 Jarnik-Prim-Dijkstra Algorithm and Related Solutions

Several cost allocation rules discussed in the literature are based on the construction of a mcst through the following algorithm, originally due to Jarnik (1930). Given $N \subset \mathcal{N}$, let $\sigma : N \rightarrow$

$\{1, 2, ..., |N|\}$ denote a priority ordering of the users in N, i.e. i is before j in the priority ordering σ if and only if $\sigma(i) < \sigma(j)$. We let $\Phi(N)$ denote the set of priority orderings of N. The Jarnik-Prim-Dijkstra algorithm (from now on J-P-D algorithm) reaches an mcst in the following manner.

Step 0 $S^0 = \{0\}$ and $g^0 = \emptyset$.

Step 1 Take the arc $(0i^1)$ such that $c_{0i^1} = \min_{i \in N} \{c_{0i}\}$. If there are several arcs satisfying this condition, choose $(0i^1)$ where $i^1 = \arg\min_{i \in N} \sigma(i)$. Let $S^1 = \{0, i^1\}$ and $g^1 = \{(0i^1)\}$.

Step $p+1$ Take an arc (ji^{p+1}) such that $c_{ji^{p+1}} = \min_{k \in S^p, l \in N \setminus S^p} \{c_{kl}\}$. If there are several arcs satisfying this condition, choose $(0i^{p+1})$ where $i^{p+1} = \arg\min_{i \in N \setminus S^p} \sigma(i)$. Let $S^{p+1} = S^p \cup \{i^{p+1}\}$ and $g^{p+1} = g^p \cup \{(ji^{p+1})\}$.

This algorithm terminates after $|N|$ steps. Note that different priority orderings may lead to different minimum cost spanning trees. We write $g^\sigma[N_0, c]$ to denote the mcst resulting from the priority ordering σ. The total cost of $g^\sigma[N_0, c]$ is the same for all $\sigma \in \Phi(N)$ (see Prim (1956)). It follows that when we have a unique mcst, $g^\sigma[N_0, c]$ is independent of σ.

Given a tree g, the immediate predecessor of user i, denoted by $\alpha^{(N,c)}(i, g)$ is the user that comes right before i : $(\alpha^{(N,c)}(i, g), i) \in g$ and $\alpha^{(N,c)}(i, g)$ is on the (unique) path from 0 to i. The successors of user i are those users that come after i : $S^{(N,c)}(i, g) = \{k \in N |\ i \text{ is on the path from 0 to } k\}$. The immediate successors of user i is denoted by $\beta^{(N,c)}(i, g) = \{k \in N | \alpha^{(N,c)}(k, g) = i\}$. When there is no room for confusion, we will drop the superscript (N, c) and the argument of g.

Given a priority ordering σ, Bird's allocation on the mcst $g^\sigma[N, c]$ is given by

$$B_i^\sigma(N, c) = c_{\alpha(i, g^\sigma[N,c])i} \ \forall i \in N.$$

When there are several mcst, we can extend Bird's rule by taking a weighted average over the priority orderings according to which ties are broken in the J-P-D algorithm, i.e.

$$B(N, c) = \sum_{\sigma \in \Phi(N)} w_\sigma B^\sigma(N, c)$$

where $w_\sigma > 0$ for all $\sigma \in \Phi(N)$ and $\sum_{\sigma \in \Phi(N)} w_\sigma = 1$. A natural extension is by setting $w_\sigma = \frac{1}{|N|!}$ for all $\sigma \in \Phi(N)$. When there is no room for confusion, we write B_i instead of $B_i(N, c)$. Also, we use $B_i^\sigma(N, c)$ to denote $c_{\alpha(i, g^\sigma[N_0, c])i}$ for notational ease.

For the sake of completeness, we give a recursive definition of the Dutta-Kar rule, which is different from, yet equivalent to, the definition that appears in Dutta and Kar (2004). Given a mcst problem (N, c), $|N| = n$, $c \in C_N^1$, let $\{1, 2, ..., n\}$ be the order induced by the J-P-D algrorithm, i.e. 1 is the first to connect to the network, 2 is second, and so forth. For each $i \in N$, let $e_i = (\alpha(i)\, i)$. Let $m = \arg\max_{i \in N} c_{e_i}$. The DK rule charges the cost of the most expensive link in the mcst, c_m, to user n,i.e. $DK_n = c_{e_m}$. For all users $i > m, i \neq n$, we have $DK_m = c_{e_{m+1}}$. Applying this procedure to the mcst problem $\left(\{1, 2, ..., m-1\}, c_{\{0,1,2,...,m-1\}}\right)$, we get the DK allocation of at least one more user. Therefore, recursive application defines the rule in a finite number of steps.

2.3 Merge-Proofness

We consider two types of strategic behavior: covert-merging and overt-merging. Under a *covert-merging maneuver* by a group of users $S \subset N$, S chooses only one user $i^* \in S$ to remain in the network. The users $S \backslash i^*$ disappear from the problem, and will get the service through user i^* in a covert fashion, by constructing a private network among themselves. Under an *overt-merging maneuver* by S, on the other hand, S announces overtly that they merged under a single identity i^*, and they demand to be treated as if they were a single user with name i^*. With both types of maneuvers, the new problem has $(N \backslash S) \cup i^*$ as its user set. When we are updating the cost matrix, the distinction between covert-merging and overt-merging maneuvers becomes relevant. Let the updated cost matrix be denoted by $c_{S \to i^*}$. Under a covert-merging maneuver, we are unaware of the existence of the users $S \backslash i^*$, i.e.

$$\begin{aligned} (c_{S \to i^*})_{i^* j} &= c_{i^* j} \text{ for all } j \in N_0 \backslash S \\ (c_{S \to i^*})_{ij} &= c_{ij} \text{ for all } (ij) \subset N_0 \backslash S \end{aligned} \tag{1}$$

If the merging is done overtly, we know that all the users in S will be in the network. Thus, when the service provider wants to connect the representative user i^* to some $j \in N \backslash S$, it can use any of the arcs $(ij)\,, i \in S$. This gives rise to the following update:

$$\begin{aligned} \left(c_{S \to i^*}\right)_{i^* j} &= \min_{i \in S} c_{ij} \text{ for all } j \in N_0 \backslash S \\ \left(c_{S \to i^*}\right)_{ij} &= c_{ij} \text{ for all } (ij) \subset N_0 \backslash S \end{aligned} \tag{2}$$

Merging maneuvers are undesirable from an efficiency viewpoint. The private network constructed by the merging coalition S, as well as the mcst constructed in the merged problem, typically involves links that would not have been part of the network had S not merged. As a result, merging increases the overall cost of the network.

When a merging maneuver occurs, the users in S construct a private network among themselves, and then they connect to the network through the merged user i^*. Accordingly, when S merges, they have to pay for the cost share of i^* in the merged problem $\left((N \backslash S) \cup i^*, c_{S \to i^*}\right)$, and they pay for the cost of $g\left[S, c_S\right]$, which we defined to be $m\left(S, c_S\right)$. We are now ready to give the following definition.

Definition 15 *A cost allocation rule ϕ is vulnerable to covert-merging (overt-merging) on $\mathcal{D} \subset \mathcal{C}$ if there exist $N \subset \mathcal{N}$,$c \in \mathcal{D}, S \subset N$, and $i^* \in S$ such that*

$$\sum_{i \in S} \phi_i\left(N, c\right) > \phi_{i^*}\left(\left(N \backslash S\right) \cup i^*, c_{S \to i^*}\right) + m\left(S, c_S\right). \tag{3}$$

where $c_{S \to i^*}$ is defined through equations 1 (equations 2).

We say that ϕ satisfies covert-merge–proofness (overt-merge–proofness) on $\mathcal{D}$ if it is not vulnerable to covert-merging (overt-merging) on $\mathcal{D}$.

The above definition of overt-merge–proofness is restrictive in the sense that the service provider is assumed to treat the merged user i^* as if he was a single identity, even though it may be aware of the fact that i^* is the representative user for coalition S (as reflected in the updated cost matrix).

For instance, the definition precludes the possibility of charging a fee to S. Overt-merge–proofness would be easy to attain through the use of such punishment mechanisms. Here, we are interested in the case where S has the legal right to become a single identity.

We abbreviate overt-merge–proofness by OMP. The following propositions show that both covert- and overt-merge–proofness are very strong properties.

Proposition 2 *On $\mathcal{C}^2$, there is no covert-merge–proof cost allocation rule.*

Proof. Take a cost allocation rule ϕ and suppose that it is covert-merge–proof. Consider the mcst problem P of Figure 1: $N = \{A, B, C\}$, $c_{0A} = 10, c_{0B} = 11, c_{0C} = 12, c_{AB} = 2, c_{AC} = 3, c_{BC} = 4$. Let $\phi = \phi(N, c)$. Note that $\phi_A + \phi_B + \phi_C = m(N_0, c) = 15$. We now show that covert-merge–proofness requires the contradicting inequality $\phi_A + \phi_B + \phi_C \leq 13$ by placing restrictions on ϕ_A, ϕ_B, and ϕ_C.

1. $\phi_A \leq 5$

covert-merge–proofness places the following restrictions on ϕ:

$$\phi_A + \phi_B \leq \phi_B\left(N\backslash A, c_{\{A,B\}\to B}\right) + 2$$

$$\phi_A + \phi_C \leq \phi_C\left(N\backslash A, c_{\{A,C\}\to C}\right) + 3$$

We have $\phi_B\left(N\backslash A, c_{\{A,B\}\to B}\right) + \phi_C\left(N\backslash A, c_{\{A,C\}\to C}\right) = 15$ since $c_{\{A,B\}\to B} = c_{\{A,C\}\to C}$. Adding these two inequalities yields $\phi_A \leq 5$.

2. $\phi_B \leq 4$

covert-merge–proofness places the following restrictions on ϕ:

$$\phi_A + \phi_B \leq \phi_A\left(N\backslash B, c_{\{A,B\}\to A}\right) + 2$$

$$\phi_B + \phi_C \leq \phi_C\left(N\backslash B, c_{\{B,C\}\to C}\right) + 4$$

We have $\phi_A\left(N\backslash B, c_{\{A,B\}\to A}\right) + \phi_C\left(N\backslash B, c_{\{B,C\}\to C}\right) = 13$, since $c_{\{A,B\}\to A} = c_{\{B,C\}\to C}$. Adding these two inequalities yields $\phi_B \leq 4$.

3. $\phi_C \leq 4$

covert-merge–proofness places the following restrictions on ϕ:

$$\phi_A + \phi_C \leq \phi_A\left(N\backslash C, c_{\{A,C\}\to A}\right) + 3$$

$$\phi_B + \phi_C \leq \phi_B\left(N\backslash C, c_{\{B,C\}\to B}\right) + 4$$

We have $\phi_A\left(N\backslash C, c_{\{A,C\}\to A}\right)+\phi_B\left(N\backslash C, c_{\{B,C\}\to B}\right)=12$, since $c_{\{A,C\}\to A}=c_{\{B,C\}\to B}$. Adding these two inequalities yields $\phi_C \leq 4$. ■

Proposition 3 *On $\mathcal{C}$, there is no overt-merge–proof cost allocation rule.*

Proof. Take a cost allocation rule ϕ and suppose that it is overt-merge–proof. Consider the following problem: $N=\{A,B,C\}$, $c_{0i}=10$ for $i=A,B,C$ and $c_{ij}=1$ for all $i,j\in N, i\neq j$. Now, the cost of a mcst is $m\left(N_0,c\right)=12$. Let $\phi=\phi\left(N,c\right)$, and without loss of generality, assume that $\phi_A\leq\phi_B\leq\phi_C$. Note that this implies $\phi_C\geq 4$.

Consider the problem (N^*,c^*) with $N^*=\{A,B\}$, $c^*_{0A}=10$ for $i=A,B$ and $c^*_{AB}=1$. Let $\phi^*=\phi\left(N^*,c^*\right)$. Note that (N^*,c^*) is the resulting problem under an overt-merging maneuver by $\{A,C\}$ into A and $\{B,C\}$ into B. OMP implies

$$\phi_A+\phi_C\leq 1+\phi^*_A$$

$$\phi_B+\phi_C\leq 1+\phi^*_B$$

Since $\phi_A+\phi_B+\phi_C=12$ and $\phi^*_A+\phi^*_B=11$, adding these two equations yields $\phi_C\leq 1$, a contradiction. ■

Since on the general domain $\mathcal{C}$ no cost allocation rule satisfies overt-merge–proofness, we confine ourselves to a smaller domain of cost matrices.

Proposition 4 *On C^1, Bird's rule satisfies OMP.*

Proof. Consider a mcst problem (N,c), $N\subset\mathcal{N}$ and $c\in C^1_N$. Take any $S\subset N$. Suppose S merges into $i^*\in S$. The new problem becomes $(N^*,c_{S\to i^*})$, where $N^*=(N\backslash S)\cup i^*$ and $c_{S\to i^*}$ is as defined above.

It is possible that $c_{S\to i^*}\notin C^1$, in which case Bird takes averages over different priority orderings of N^*. Let B^* be any extension of Bird's rule, i.e.

$$B^*\left(N^*,c_{S\to i^*}\right)=\sum_{\sigma\in\Phi(N^*)}w_\sigma B^\sigma\left(N^*,c_{S\to i^*}\right)$$

where $w_\sigma > 0$ for all $\sigma \in \Phi(N)$ and $\sum_{\sigma \in \Phi(N)} w_\sigma = 1$.

Let $B = B(N, c)$ and $B^* = B(N^*, c_{S \to i^*})$. Choose $\sigma^* \in \Phi(N^*)$ such that $B_{i^*}^{\sigma^*} = \min_{\sigma \in \Phi(N^*)} B_{i^*}^{\sigma}(N^*, c_{S \to i^*})$. Note that there exists $i^1 \in S$ such that $c_{(\alpha(i^*) i^1)} = c_{(\alpha(i^*) i^*)}$. Consider an ordering σ of N such that, for all i with $\sigma^*(i) < \sigma^*(i^*)$, $\sigma(i) = \sigma^*(i)$ and $\sigma(i^1) = \sigma^*(i^*)$.

Now let τ and τ^* denote the orderings according to which the users in N and N^* are added to the networks $g^\sigma[N_0, c]$ and $g^{\sigma^*}[N_0^*, c_{S \to i^*}]$ in the J-P-D algorithm, respectively. Until the step at which the merged user i^* is added to the network, none of the comparisons made by the J-P-D algorithm are affected by the merging maneuver. This implies that, for all users with $\tau^*(i) < \tau^*(i^*)$, we have $\tau(i) = \tau^*(i)$ and $\tau(i^1) = \tau^*(i^*)$. Hence, $B_{i^1} = B_{i^*}^{\sigma^*}$. It follows from this argument that in graph $g^\sigma[N_0, c]$, none of the users in $S \backslash i^1$ lie on the path from 0 to i^1.

We showed that $B_{i^1} = B_{i^*}^{\sigma^*} \leq B_{i^*}^*$. A sufficient condition for the merging move not to be beneficial is $\sum_{i \in S \backslash i^1} B_i \leq m(S, c_S)$.

Now let $g^1 = \{(\alpha(i)\, i)\}_{i \in (N \backslash S) \cup i^1}$ and consider the graph $g = g^1 \cup g[S, c_S]$. Note that g^1 is a subgraph of $g[N_0, c]$ and is therefore a forest. Let $G_1, G_2, ..., G_K$ be the set of disjoint trees in g^1 : $G_k \cap G_l = \emptyset$ for all $k, l = 1, 2, ..., K, k \neq l$, and $\cup_{k=1}^K G_k = g^1$. Without loss of generality, assume that $i^1 \in G_1$. This implies that $0 \in G_1$ as well, since we argued above that none of the users in $S \backslash i^1$ lie on the path from 0 to i^1. Now, as all users in $S \backslash i^1$ are connected to i^1 through the network $g[S, c_S]$, they are also connected to 0 in g. By construction, for all $k = 2, 3, ..., K$ there exists a user j_k in tree G_k such that $\alpha(j_k) \in S \backslash i^1$. This in turn implies that all users in $\cup_{k=2}^K G_k$ are connected to 0 in graph g.

Hence, g is a connected graph. As g^1 has $|N| - |S| + 1$ edges and $g[S, c_S]$ has $|S| - 1$ edges, g is a connected graph with $|N| + 1$ vertices and $|N|$ edges. This implies that g must be a tree, and the cost of g must be no less than the cost of the mcst on N_0, i.e.

$$\sum_{i \in (N \backslash S) \cup i^1} B_i + m(S, c_S) \leq \sum_{i \in (N \backslash S) \cup i^1} B_i + \sum_{i \in S \backslash i^1} B_i.$$

This immediately gives us the desired inequality, establishing that B is overt-merge–proof. ■

The literature on mcst usually assumes that any group of users has access to the network technol-

ogy. In other words, a group of users $S \subset N$ can break away from the grand coalition and construct their own network to connect to the source. We assume that they can only do this using their own links.[11] Accordingly, the worth of coalition S is defined to be the cost of a mcst on $S \cup \{0\}$, i.e. $m(S_0, c_{S_0})$. The notion of the core of an mcst game thus incorporates two normative properties: stability (no coalition breaks away from the system) and fairness (no coalition subsidizes another). We say that a cost allocation $z \in \mathbb{R}^N$ is in the core of the mcst problem (N, c) if, $\sum_{i \in N} z_i = m(N_0, c)$ and for all $S \subset N$, we have $\sum_{i \in S} z_i \leq m(S_0, c_{S_0})$. We write $Core(N, c)$ to denote the set of cost allocations that are in the core of (N, c).

Definition 16 *A cost allocation rule ϕ satisfies core selection (CS) on $\mathcal{D} \subset \mathcal{C}$ if, for all $N \subset \mathcal{N}, c \in \mathcal{D}$, $\phi(N, c) \in Core(N, c)$.*

The following property for cost allocation rules suggests that we use only the information from the mcst while allocating costs.

Definition 17 *A cost allocation rule ϕ satisfies tree invariance (TI) on $\mathcal{D} \subset \mathcal{C}$ if, for all $N \subset \mathcal{N}$ and $c, \bar{c} \in \mathcal{D}$ such that c and $\bar{c}$ have the same mcst*

$$\phi_i(N, c) = \phi_i(N, \bar{c}) \text{ for all } i \in N.$$

Tree invariance is partially justified by the computational complexity associated with the rules not satisfying this property. We are now ready to state our main result.

Theorem 3 *On $\mathcal{C}^2$, the cost allocation rule ϕ satisfies OMP, CS, and TI if and only if $\phi = B$.*

Proof. We already know B satisfies CS (Granot and Huberman, 1981). Also, Bird's rule is defined using information only from the mcst of a mcst problem, so that it also satisfies TI. We showed that B satisfies OMP in Proposition 3.

To show the only if part, take a cost allocation rule ϕ that satisfies OMP, CS, and TI. We use an induction argument on $|N|$ to show that $\phi = B$. It is clear that ϕ coincides with B on problems

[11] See footnote 8.

with $|N| = 1$.

Induction Argument 1: Assume that ϕ coincides with B on problems with $|N| \leq K$, where $1 \leq K$.

Now consider a mcst problem (N, c) with $|N| = K + 1$, and let $\phi = \phi(N, c)$.

If all users are leaves in $g[N_0, c]$, CS implies $\phi_i \leq B_i$ for all $i \in N$. Since $\sum_{i \in N} \phi_i = \sum_{i \in N} B_i$, we have $\phi_i = B_i$ for all $i \in N$.

From now on, assume that there exists a user $i \in N$ who is not a leaf in $g[N_0, c]$. We call i a non-leaf user. We use an induction argument on the number of users in $S(i)$ for a non-leaf user i to show that $\phi = B$. We make use of the following observations.

Observation 1 *$\sum_{j \in i \cup S(i)} \phi_j = \sum_{j \in i \cup S(i)} B_j$ for all non-leaf users $i \in N$.*

Proof. Consider a non-leaf user $i \in N$. Since the worth of the coalition $N \setminus (i \cup S(i))$ is $\sum_{j \in N \setminus (i \cup S(i))} B_j$, CS implies $\sum_{j \in i \cup S(i)} \phi_j \geq \sum_{j \in i \cup S(i)} B_j$.

When the coalition $i \cup S(i)$ overt-merges into i, the new problem is one with less than K players and hence by the Induction Argument 1, the merged user pays B_i in the new problem. As the cost of the spanning tree on $i \cup S(i)$ is $\sum_{j \in S(i)} B_j$, OMP implies $\sum_{j \in i \cup S(i)} \phi_j \leq \sum_{j \in i \cup S(i)} B_j$.

■

Observation 2 *Consider $i_1, i_2 \in N$ such that $(i_1 i_2) \notin g[N_0, c]$ and $\alpha(i_1) = \alpha(i_2)$, and the problem $(N, \bar{c})$ where $\bar{c}_{ij} = c_{ij}$ for $\{i, j\} \neq \{i_1, i_2\}$ and $\bar{c}_{i_1 i_2} > \max\{B_{i_1}, B_{i_2}\}$. We have $g[N_0, c] = g[N_0, \bar{c}]$.*

Proof. This follows directly from the construction of a mcst using the J-P-D algorithm. ■

Observation 3 For a non-leaf user i such that $|S(i)| = 1$, $\phi_i = B_i$.

Proof. Let $S(i) = i_1$, $a = B_i$ and $b = B_{i_1}$. Observation 1 implies $\phi_i + \phi_{i_1} = a + b$.

As i_1 is a leaf, $S(i_1) = \emptyset$ and CS implies that $\phi_{i_1} = b + x$ for some $x \geq 0$. Suppose $x > 0$. Take a user $i_2 \notin N$, and consider the problem $(\bar{N}, \bar{c})$, where $\bar{N} = N \cup \{i_2\} \setminus \{i_1\}$, $\bar{c}_{ij} = c_{ij}$ for all $\{i, j\} \subset \bar{N} \setminus \{i_2\}$, $\bar{c}_{ii_2} = d < b$, and $\bar{c}_{ji_2} > m(N_0, c)$ for all $j \in \bar{N} \setminus \{i\}$. Check that

$\beta^{(\bar{N},\bar{c})}(i) = i_2$. Let $\bar{\phi} = \phi(\bar{N},\bar{c})$. CS implies $\bar{\phi}_{i_2} = d + y$ for some $y \geq 0$, and Observation 1 implies $\bar{\phi}_i + \bar{\phi}_{i_2} = a + d$.

Next, consider the problem $\left(\tilde{N},\tilde{c}\right)$, where $\tilde{N} = N \cup \{i_2\}$, $\tilde{c}_{ij} = c_{ij}$ for all $\{i,j\} \subset \tilde{N} \backslash \{i_2\}$, $\tilde{c}_{ji_2} = \bar{c}_{ji_2}$ for all $j \in \tilde{N} \backslash \{i_1\}$ and $c_{i_1 i_2} = b + \varepsilon$ for some $0 < \varepsilon < x$. Check that $\beta^{(\tilde{N},\tilde{c})}(i) = \{i_1, i_2\}$ and $\{i_1, i_2\}$ are leaves in graph $g\left[\tilde{N}_0, \tilde{c}\right]$.

Let $\tilde{\phi} = \phi\left(\tilde{N},\tilde{C}\right)$ and note that CS implies $\sum_{j\in\{i,i_1,i_2\}} \tilde{\phi}_j \geq \sum_{j\in\{i,i_1,i_2\}} B_j$. Also, after an overt-merging maneuver by $\{i, i_1, i_2\}$ into i, the new problem has $N-1$ users and by Induction Argument 1, everyone pays their Bird allocations. As this maneuver should be unprofitable, we must have $\sum_{j\in\{i,i_1,i_2\}} \tilde{\phi}_j \leq \sum_{j\in\{i,i_1,i_2\}} B_j$ and therefore we have $\tilde{\phi}_i + \tilde{\phi}_{i_1} + \tilde{\phi}_{i_2} = a + b + d$. Now, suppose $\tilde{\phi}_{i_1} < b + x$, which implies $\tilde{\phi}_i + \tilde{\phi}_{i_2} > a - x + d$. Consider an overt-merging move by users $\{i, i_2\}$ into $\{i\}$.The mcst corresponding to the resulting problem is the same as that of the original problem (N, c) and so by TI, user $\{i\}$ pays $a - x$ in the merged problem. Hence, together with the cost of the private network, $\{i, i_2\}$ coalition pays $a - x + d$ when they merge, a contradiction to OMP. Hence, $\tilde{\phi}_{i_1} \geq b + x$.

A similar argument to the one above shows $\tilde{\phi}_{i_2} \geq d + y$. Now, suppose $\{i_1, i_2\}$ coalition overt-merges into i_2. Again, the mcst corresponding to the resulting problem is the same as that of the problem $(\bar{N},\bar{c})$ and so by TI, user i_2 pays $d + y$ in the merged problem. The cost of the private network $(i_1 i_2)$ is $b + \varepsilon$. Hence, after the overt-merging maneuver, coalition $\{i_1, i_2\}$ pays $d + y + b + \varepsilon < d + y + b + x \leq \tilde{\phi}_{i_1} + \tilde{\phi}_{i_2}$. This is a contradiction to OMP, establishing that $\phi_{i_1} = b$, and since $\phi_i + \phi_{i_1} = a + b$, $\phi_i = a$. ■

Induction Argument 2: Fix a positive integer $M \leq K - 1$ and assume that, for all users $j \in N$ such that $S(j)$ contains at most M users, $\phi_k = B_k$ for all $k \in j \cup S(j)$.

Consider a user $i \in N$ such that $S(i)$ contains $M + 1$ users. We consider three cases.

Case 1. *Every user in $\beta(i)$ is a non-leaf user.* Let $\beta(i) = \{i_1, i_2, \cdots, i_m\}$. Now, by Induction

Argument 2, for all $l = 1, 2, ..., m$, we have $\phi_j = B_j$ for all $j \in i_l \cup S(i_l)$. Therefore,

$$\sum_{j \in S(i)} \phi_j = \sum_{l=1}^{m} \sum_{j \in i_l \cup S(i_l)} \phi_j = \sum_{l=1}^{m} \sum_{j \in i_l \cup S(i_l)} B_j = \sum_{j \in S(i)} B_j$$

Since $\sum_{j \in i \cup S(i)} \phi_j = \sum_{j \in i \cup S(i)} B_j$ by Observation 1, we have $\phi_i = B_i$.

Case 2. *Every user in $\beta(i)$ is a leaf.* Let $\beta(i) = \{i_1, i_2, ..., i_m\}, m \geq 2$.

Let $a = B_i$ and assume, without loss of generality, that $B_{i_1} \leq B_{i_2} \leq \cdots \leq B_{i_m}$. For each $i_k \in \beta(i)$, CS implies $\phi_{i_k} = B_{i_k} + x_k$ for some $x_k \geq 0$. Now, suppose there exists $k \in \{1, 2, ..., m\}$ such that $x_k > 0$. Take any $i_l \neq i_k$ and consider the problem $(N, \bar{c})$ where $\bar{c}_{ij} = c_{ij}$ for all $\{i, j\} \neq \{i_k, i_l\}$, and $\bar{c}_{i_k i_l} = \max\{B_{i_k}, B_{i_l}\} + \varepsilon$ for some $0 < \varepsilon < x_k$. By Observation 2, the mcst corresponding to the problem $(N, \bar{c})$ is the same as that of (N, c). This implies that $\phi(N, \bar{c}) = \phi(N, c) = \phi$. When the coalition $\{i_k, i_l\}$ merges into $i_{\arg\min\{k,l\}}$, they have to construct the edge $(i_k i_l)$ which costs $\max\{B_{i_k}, B_{i_l}\} + \varepsilon$. The merged user pays his Bird allocation $\arg\min\{B_{i_k}, B_{i_l}\}$ by Induction Argument 1. Now OMP implies $\min\{B_{i_k}, B_{i_l}\} + \max\{B_{i_k}, B_{i_l}\} + \varepsilon = B_{i_k} + B_{i_l} + \varepsilon \geq \phi_1 + \phi_k = B_{i_k} + B_{i_l} + x_k$, a contradiction. Hence, $x_k = 0$ for all $k \in \{1, 2, ..., m\}$. This, together with Observation 1, implies that $\phi_i = B_i$.

Case 3. $\beta(i) = \{i_1, i_2, \cdots, i_m, j_1, j_2, \cdots, j_n\}$, *where every user in $\{i_1, i_2, \cdots, i_m\}$ is a non-leaf user and every user in $\{j_1, j_2, \cdots, j_n\}$ is a leaf.* For all $l = 1, 2, ...m$, $\beta(i_l)$ contains at most M non-leaf users, and hence by Induction Argument 2, $\phi_j = B_j$ for all $j \in i_l \cup S(i_l)$. For every $l = 1, 2, ..., n$, CS implies $\phi_{j_l} = B_{j_l} + x_l$ for some $x_l \geq 0$. Now suppose there exist $j_l \in \{j_1, j_2, \cdots, j_n\}$ such that $\phi_{j_l} = B_{j_l} + x_l$ for some $x_l > 0$. Next, consider the problem $(N, \bar{c})$ where $\bar{c}_{ij} = c_{ij}$ for all $\{i, j\} \neq \{i_1, j_l\}$ and $\bar{c}_{i_1 j_l} = \max\{B_{i_1}, B_{j_l}\} + \varepsilon$ for some $0 < \varepsilon < x_l$. We know that the mcst corresponding to $(N, \bar{c})$ is the same as that of (N, c) by Observation 2, and so by TI, $\phi(N, \bar{c}) = \phi(N, c) = \phi$. Consider an overt-merging maneuver by $S = \{i_1, j_l\}$ into $i^* = \arg\min_{j \in S} B_j$. The new problem has at most K users and so by Induction Argument 1, we have $\phi_{i^*}(N, \bar{c}) = B_{i^*}$. This means that in problem $(N, \bar{c})$, the cost share to coalition S, $B_{i_1} + B_{j_l} + x_l$, is more than what they would have paid had they merged, $\min\{B_{i_1}, B_{j_l}\} + \max\{B_{i_1}, B_{j_l}\} + \varepsilon = B_{i_1} + B_{j_l} + \varepsilon$. This is a contradiction to OMP, and

therefore we must have $x_{j_l} = 0$ for all $l \in \{j_1, j_2, \cdots, j_n\}$. In other words, $\phi_j = B_j$ for all $j \in S(i)$. This, together with Observation 1, implies that $\phi_i = B_i$.

This ends the induction argument, and establishes that $\phi = B$. ■

2.4 Discussion

2.4.1 Implications of the main result

We first introduce the notion of the *irreducible core* of a mcst problem (N, c), denoted by $IrrCore(N, c)$, first discussed in Bird (1976). An allocation x is in $IrrCore(N, c)$ if, for all cost matrices $c' \in C_N$ such that $g[N_0, c]$ is a mcst in problem (N_0, c'), we have $x \in Core(N, c')$. Following Bird (1976), given a mcst problem $(N, c), c \in C_N^1$, define the *canonical cost matrix* of (N, c), denoted by $\tilde{c}(N, c)$, as follows.[12]

i) If $(ij) \in g[N_0, c]$, then $[\tilde{c}(N, c)]_{ij} = c_{ij}$, and

ii) if $(ij) \notin g[N_0, c]$, then $[\tilde{c}(N, c)]_{ij} = \max c_{kl}$ where $(kl) \in g[N_0, c]$ and is in the cycle formed by the addition of (ij) to the mcst $g[N_0, c]$.

Bergantiños and Vidal-Puga (2004) prove that $\tilde{c}(N, c)$ is unique. For an example, consider the problem $(N, \bar{c})$ where $N = \{A, B, C\}$, $\bar{c}_{0A} = 10, \bar{c}_{AB} = \bar{c}_{AC} = 1$, and $\bar{c}_{0B} = \bar{c}_{0C} = 15, \bar{c}_{BC} = 5$. The canonical cost matrix $\tilde{c}(N, \bar{c})$ is given by the cost matrix c in the proof of Proposition 2.

A cost allocation rule ϕ is a selection from the irreducible core on $\mathcal{D} \subset \mathcal{C}$ if, for all $N \subset \mathcal{N}, c \in \mathcal{D}$ we have $\phi(N, c) \in Core(N, \tilde{c}(N, c))$.[13] Feltkamp et. al. (1994) constructs a class of cost allocation rules, which they call *obligation rules*, and are selections from the irreducible core. Indeed, for every allocation in the irreducible core, there exists an obligation rule that leads to that allocation. They propose the *equal remaining obligations rule* as a selection from this class. Also among these rules are Bird's rule and the Dutta-Kar rule. As mentioned in that paper, $Core(N, \tilde{c}(N, c)) \subset Core(N, c)$.

[12] Bird (1976) calls $\tilde{c}(N, c)$ the *minimal network*.

[13] The cost sharing game $(N, \tilde{c}(N, c))$ is submodular (Bird, 1976), and the extreme points of the irreducible core are the Bird allocations induced by all the different priority orderings of N. As a result, the *Shapley value* of $(N, \tilde{c}(N, c))$ coincides with Bird's rule as it is applied to the problem $(N, \tilde{c}(N, c))$. This is also the same rule as the equal remaining obligations rule, and the *p-value* as it appears in Branzei, Moretti, Norde and Tijs (2003). See these papers and Bergantiños and Vidal-Puga (2004) for alternative characterizations of this rule.

Therefore, if the cost allocation ϕ is an irreducible core selection, it satisfies CS. Also, since these rules use only information from the mcst, they satisfy TI. Our main result implies that, on domain $\mathcal{C}^2$, Bird's rule is the only overt-merge–proof cost allocation rule among these rules.

Besides the one due to Gómez-Rúa and Vidal-Puga (2006), two other characterizations of Bird' rule appear in the literature, and they both rely on consistency properties. Feltkamp et. al. (2000) uses *consistency* and *converse consistency.* These properties relate the cost allocation of a mcst problem (N, c) to the cost allocation of its reduced problem, which is obtained by deleting a leaf user from the problem. The consistency property used in Dutta and Kar's characterization of Bird's rule, on the other hand, looks at the reduced problem obtained by deleting a user i who is adjacent to the source. Suppose user i walks away from the problem with his cost share ϕ_i. One can argue that, in the reduced problem, the cost of connecting $j \neq i$ to the source becomes $\min\{c_{0j}, c_{0i} + c_{ij} - \phi_i\}$. *Source consistency* requires that the cost allocation stays the same in the reduced problem, and characterizes Bird's rule together with the property of *extreme point monotonicity,* which asserts that deleting a leaf from the problem should lead to a higher cost share for all users.

Feltkamp et. al. (2000) looks at the Nash equilibria of a strategic form game, where each player i chooses, for each $j \in N_0 \backslash i$, how much to contribute to the cost of the edge (ij). Recall that given $\sigma \in \Phi(N)$, B^σ denotes the Bird allocation as it is applied to the tree $g^\sigma[N_0, c]$. The set of Nash equilibria of this game contains B^σ for all $\sigma \in \Phi(N)$. However, some of these equilibria may correspond to the construction of an inefficient network.

2.4.2 Independence of the axioms

We provide examples to illustrate the tightness of the result in Theorem 1. The Dutta-Kar rule satisfies CS and TI, but is not OMP as demonstrated in the introduction.

Given problem (N, c), let i_1 be the first user that connects to the network in the J-P-D algorithm and define $\tilde{\phi}_{i_1}(N, c) = m(N_0, c)$, $\tilde{\phi}_i(N_0, c) = 0$ for all $i \neq i_1$. Check that $\tilde{\phi}$ satisfies TI but not CS. To see that it satisfies OMP, first note that any beneficial overt-merging maneuver has to include i_1. For any S such that $i_1 \in S$, when S overt-merges into $i^* \in S$, the merged user i^* pays the cost of the

network $g\left[\left(\left(N\backslash S\right)\cup i^{*}\right)_{0},c_{S\rightarrow i_{1}}\right]$. S also pays the cost of the private network $g\left[S,c_{S}\right]$. Note that the graph $g=g\left[\left(\left(N\backslash S\right)\cup i^{*}\right)_{0},c_{S\rightarrow i_{1}}\right]\cup g\left[S,c_{S}\right]$ is connected. Therefore, the total cost to S after the overt-merging maneuver is the cost of graph g, which cannot be less than $m\left(N_{0},c\right)=\sum_{i\in S}\bar{\phi}_{i}\left(N,c\right)$.

We now give an example of a rule $\bar{\phi}$ that satisfies CS, OMP, but not TI. The rule $\bar{\phi}$ will be the same as Bird's rule, except when there exists a user with very limited merging capabilities. Consider a leaf i such that c_{ij} is very high for all $j\neq\alpha\left(i\right)$, so that user i cannot profitably merge with users beside $\alpha\left(i\right)$. Let $\bar{\phi}_{i}=B_{i}+x$ for some $x>0$, but give the residual x to $\alpha\left(i\right):\bar{\phi}_{\alpha(i)}=B_{\alpha(i)}-x$. This ensures that i cannot profitably merge with $\alpha\left(i\right)$ either. Let $\bar{\phi}_{i}=B_{i}$ for all $i\notin\left\{i,\alpha\left(i\right)\right\}$ so that any coalition that does not include i cannot merge profitably either. We now introduce $\bar{\phi}$ formally and prove its properties.

Given problem $\left(N,c\right),\left|N\right|=n$, set $\delta\left(N,c\right)=\min\left\{c_{ij}\text{ s.t. }i,j\in N_{0},i\neq j\right\}$. Let i_{n} be the last user that connects to the network in the J-P-D algorithm. Note that $B_{i_{n}}$ is the cost of user i_{n}'s cheapest link. Let $\Delta\left(N,c\right)$ be the cost of user i_{n}'s second cheapest link. If $\Delta\left(N,c\right)>m\left(N_{0},c\right)$ and $\nexists i\in N$ such that $\max_{j\in N_{0}\backslash i}c_{ij}>m\left(N_{0},c\right)$, let $\bar{\phi}_{i_{n}}\left(N,c\right)=B_{i_{n}}+\delta$, $\bar{\phi}_{\alpha(i_{n})}\left(N,c\right)=B_{\alpha(i_{n})}-\delta$, and $\bar{\phi}_{i}\left(N,c\right)=B_{i}$ for all $i\notin\left\{i_{n},\alpha\left(i_{n}\right)\right\}$. Otherwise, let $\bar{\phi}\left(N,c\right)=B\left(N,c\right)$. Clearly, $\bar{\phi}$ does not satisfy TI.

We first show that $\bar{\phi}$ satisfies OMP. Fix a problem $\left(N,c\right)$, and let $\bar{\phi}=\bar{\phi}\left(N,c\right)$, $B=B\left(N,c\right)$. Consider an overt-merging maneuver by a coalition $S\subset N$ into $i^{*}\in S$ and let $\bar{\phi}^{*}=\bar{\phi}\left(\left(N\backslash S\right)\cup i^{*},c_{S\rightarrow i^{*}}\right),B^{*}=B\left(\left(N\backslash S\right)\cup i^{*},c_{S\rightarrow i^{*}}\right)$. Note that, if $\bar{\phi}=B$ we also have $\bar{\phi}^{*}=B^{*}$ and the merging is not beneficial. Assume that $\bar{\phi}\neq B$, i.e. that $\Delta\left(N,c\right)>m\left(N_{0},c\right)$ and $\nexists i\in N$ such that $\max_{j\in N_{0}\backslash i}c_{ij}>m\left(N_{0},c\right)$.

Suppose that $i_{n}\in S$. Observe that, due to the updating in the cost matrix, this implies that $\Delta\left(\left(N\backslash S\right)\cup i^{*},c_{S\rightarrow i^{*}}\right)\leq m\left(N_{0},c\right)$ and therefore $\bar{\phi}^{*}=B^{*}$. If $\alpha\left(i_{n}\right)\in S$, we have $\sum_{i\in S}\bar{\phi}_{i}=\sum_{i\in S}B_{i}\leq B_{i^{*}}^{*}+m\left(S,c_{S}\right)=\bar{\phi}_{i^{*}}^{*}+m\left(S,c_{S}\right)$. If $\alpha\left(i_{n}\right)\notin S$, we have $\sum_{i\in S}\bar{\phi}_{i}<m\left(N_{0},c\right)<\Delta\left(N,c\right)<m\left(S,c_{S}\right)<\bar{\phi}_{i^{*}}^{*}+m\left(S,c_{S}\right)$.

Suppose that $i_{n}\notin S$. If $\alpha\left(i_{n}\right)\in S$, then the link $\left(i^{*}i_{n}\right)$ will be part of the mcst on $\left(N\backslash S\right)\cup i^{*}$, which implies that $\bar{\phi}_{i^{*}}^{*}=B_{i^{*}}^{*}-\delta\left(\left(N\backslash S\right)\cup i^{*},c_{S\rightarrow i^{*}}\right)=B_{i^{*}}^{*}-\delta\left(N,c\right)$. Consequently, $\sum_{i\in S}\bar{\phi}_{i}=$

$\sum_{i\in S} B_i - \delta \leq B^*_{i^*} - \delta + m(S, c_S) = \bar{\phi}^*_{i^*} + m(S, c_S)$. If $\alpha(i_n) \notin S$, then $\bar{\phi}_{i^*} = B^*_{i^*}$ since the link $(i^* i_n)$ is too costly to include in the network. This implies that $\sum_{i\in S} \bar{\phi}_i = \sum_{i\in S} B_i \leq B^*_{i^*} + m(S, c_S) = \bar{\phi}^*_{i^*} + m(S, c_S)$.

Next, we show that $\bar{\phi}$ satisfies CS. To this purpose, consider $S \subset N$. If $i_n \notin S$, then $\sum_{i\in S} \bar{\phi}_i \leq \sum_{i\in S} B_i \leq m(S_0, c_{S_0})$. Assume that $i_n \in S$. If $\alpha(i_n) \in S$, then $\sum_{i\in S} \bar{\phi}_i = \sum_{i\in S} B_i \leq m(S_0, c_{S_0})$. If $\alpha(i_n) \notin S$, then $\sum_{i\in S} \bar{\phi}_i < m(N_0, c) < m(S_0, c_{S_0})$.

The construction of $\bar{\phi}$ suggests that finding rules that satisfy OMP but not TI may involve complicated computational tasks which include assessing the merging capabilities of each user.

3 Purely Subjective Probabilistic Sophistication with Stochastic Monotonicity[14]

3.1 Introduction

The subject of this chapter is decision making under uncertainty in a purely subjective setting. The term subjective refers to the fact that the decision maker does not ascribe exogoneously given probabilities to the set of states that describe the uncertainty in the problem. We will follow the Savage (1954) approach to modeling subjective uncertainty. To fix ideas in an example, suppose you are about to bet on the outcome of the next presidential election, and you view the set of *states* to be $S = \{R, D, G\}$, where R, D, and G denote the states in which Republicans, Democrats, and the Greens win, respectively. Let the set of *outcomes*, X, be the set of monetary prizes in dollars, i.e. let $X = \mathbb{R}$. An *act* is a function that assigns an outcome to each of the states. For instance, under the act [100 on R; -80 on D; 0 on G], you receive \$100 if Republicans win, lose \$80 if Democrats win, and receive nothing if the Greens win. We use the symbols f, g to denote acts. A standard assumption is that you are able to compare any two acts, i.e. that your *preferences* over acts is *complete*. These preferences incorporate two types of information. First, they carry information about how likely you think each event (set of states) is; hence they carry information about your *beliefs*. However, two decision makers who have the exact same beliefs may have different preferences due to differing *risk attitudes,* which is the second type of information embedded in your preferences.

In Machina and Schmeidler's rendering of Savage's theory, six conditions on a decision maker's preferences allow us to put structure on both her beliefs and risk attitudes.[15] Under these conditions, there exists a unique subjective probability π on events and a utility function over the set of outcomes such that, the individual prefers act f to act g if and only if the *expected utility* under f is higher than that under g. The expected utility under an act f is the sum of the utilities of the outcomes in the range of f, calibrated by the probability of the event in which f yields that specific outcome.

[14] **This chapter is joint work with Simon Grant and Ben Polak.**

[15] Savage (1954) also needs a richer state space than in our example to derive his representation result.

The Allais paradox (1953) demonstrated that most people neither conform to Savage's second postulate (Sure-Thing Principle) nor to the expected utility hypothesis. Accordingly, when Machina and Schmeidler (1992) posed the following question, it must have appeared as a natural one to most of the audience: "What does it take for choice behavior that does not necessarily conform to the expected utility hypothesis to nonetheless be based on probabilistic beliefs?"[16]

What does it mean for preferences to be "based on probabilistic beliefs"? We use a slightly different definition from Machina and Schmedler, due to Grant (1995). Suppose that the individual's beliefs are given by π. Let $\mathbf{R}_f^\pi$ denote the *lottery* that is associated with the act f, in the following manner: To each outcome x in X, $\mathbf{R}_f^\pi$ assigns the probability under π of the event in which f yields the outcome x. We say that preferences are based on beliefs, or *probabilistically sophisticated*, when we can find a unique probability measure π such that, the decision maker is indifferent between two acts f and g whenever they yield the same lottery over outcomes, i.e. $\mathbf{R}_f^\pi$ and $\mathbf{R}_g^\pi$.

Machina and Schmeidler (1992) show that, imposing Savage's set of axioms minus **P2**, while strengthening his axiom **P4** (Weak Comparative Probability) guarantees that preferences are probabilistically sophisticated. They motivate the need to dispense with **P2** by noting that "probabilistically sophisticated non-expected utility preferences," e.g. Allais-type preferences, do not satisfy **P2**. However, such preferences satisfy the *Two-Outcome Sure-Thing Principle* (**P2****), a weaker version of **P2** that we introduce here and interpret as a notion of stochastic monotonicity for acts. We show that the conjuction of **P2**** and **P4** is equivalent to Machina and Schmeidler's stronger axiom **P4***. Hence, relying on their result enables us to show that, imposing the Savage axioms while weakening **P2** to **P2**** guarantees probabilistic sophistication. We believe the advantage of our result is that our preference conditions are simpler than those of Machina and Schmeidler.

Machina and Schmeidler (1995) characterizes probabilistically sophisticated preferences in an Anscombe-Aumann (1963) setting, by replacing the expected utility hypothesis with a new Replacement Axiom. Grant and Polak (2006) shows that this latter axiom is loosely analogous to Savage's **P4**, and obtain probabilistic sophistication using a weaker version of **P4**, in conjunction with a

[16] See Machina and Schmeidler (1992) for a more extensive discussion of the motivation behind this question.

stochastic monotonicity axiom. The current chapter suggests that results very similar to Grant and Polak (2006) can be obtained in a Savage setting.

Section 2 introduces the notation, Machina and Schmeidler's rendering of Savage's theory, and their representation result. In Section 3 we introduce our alternative axiom and representation result.

3.2 Preliminaries

Set Up.

Let $\mathcal{X}$ be an arbitrary set of outcomes (finite or infinite). Assume that uncertainty is described by a set of states, $\mathcal{S}$. For any event $E \subset S$, let E^c denote its complement in $\mathcal{S}$.

Let $\mathcal{F}$ denote the set of functions from $\mathcal{S}$ to $\mathcal{X}$ with finite range, i.e. $\mathcal{F}$ is the set of Savage acts. Furthermore, for any pair of acts f and g in $\mathcal{F}$ and any event $E \subset \mathcal{S}$, $h = [f \text{ on } E;\ g \text{ on } E^c]$ denotes the act in $\mathcal{F}$ formed from 'splicing' and 'recombining' the two acts f and g in such a way that $h(s)$ equals $f(s)$ if $s \in E$, and equals $g(s)$ if $s \in E^c$.

Let $\succsim$ denote the individual's weak preference relation on $\mathcal{F}$ (as is standard $\succ$ will denote strict preference and $\sim$ will denote indifference.) With slight abuse of notation, x denotes the constant act h, where $h(s) = x$ for all $s \in \mathcal{S}$. Hence we will write $x \succsim y$ if and only if $[x \text{ on } \mathcal{S}] \succsim [y \text{ on } \mathcal{S}]$.

Denote by $\mathcal{L}$ the set of probability measures on $\mathcal{X}$ with finite supports. We will refer to the elements of $\mathcal{L}$ as roulette lotteries or just as lotteries. For each x in $\mathcal{X}$, let δ_x denote the (degenerate) lottery that yields the outcome x with probability one. Any lottery $\mathbf{R}$ in $\mathcal{L}$ may be expressed as a probability weighted mixture of degenerate lotteries, corresponding to the outcomes in its support: that is, $\mathbf{R} = \sum_{i=1}^{m} p_i \delta_{x_i}$, where $p_i = \mathbf{R}(x_i)$ for each $i = 1, \ldots, m$.

Let $\geq^1$ (respectively, $>^1$) denote the partial ordering over roulette lotteries of (respectively, strict) First-Order Stochastic Dominance derived from $\succsim$. It is defined as follows: for any pair of lotteries $\mathbf{R} = \sum_{i=1}^{m} p_i \delta_{x_i}$ and $\mathbf{R}^* = \sum_{j=1}^{m^*} q_j \delta_{y_j}$, $\mathbf{R} \geq^1 \mathbf{R}^*$, if

$$\sum_{\{i: x_i \succsim z\}} p_i \geq \sum_{\{j: y_j \succsim z\}} q_j \text{ for all } z \in \mathcal{X}.$$

And $\mathbf{R} >^1 \mathbf{R}^*$, if, in addition, strict inequality holds for some $z \in \mathcal{X}$.

We shall refer to a preference relation defined on $\mathcal{L}$ and denoted by $\succ_L$ as *risk preferences.* We shall say the risk preferences satisfy Stochastic Monotonicity (with respect to $\succsim$) if they respect the partial ordering of First Order Stochastic Dominance, that is, $\mathbf{R} \geq^1 \mathbf{R}^*$ implies $\mathbf{R} \succsim_L \mathbf{R}^*$ and $\mathbf{R} >^1 \mathbf{R}^*$ implies $\mathbf{R} \succ_L \mathbf{R}^*$. We say that a function $V : \mathcal{L} \to \mathbb{R}$ is Mixture Continuous and Stochastically Monotonic if it is continuous in probability mixtures and $\mathbf{R} >^1 (\geq^1) \mathbf{R}^*$ implies $V(\mathbf{R}) > (\geq) V(\mathbf{R}^*)$. Furthermore, we say it is affine if for any pair of roulette lotteries $\mathbf{R}$ and $\mathbf{R}^*$ in $\mathcal{L}$, and any α in $(0,1)$, $V(\alpha\mathbf{R} + (1-\alpha)\mathbf{R}^*) = \alpha V(\mathbf{R}) + (1-\alpha) V(\mathbf{R}^*)$.

Subjective Expected Utility

A preference relation $\succsim$ on $\mathcal{F}$ is said to admit a Subjective Expected Utility representation, if there exists a unique subjective probability measure $\pi(.)$ over events and a utility function $U(.)$ on $\mathcal{L}$, unique up to positive affine transformations, such that, for all pairs of acts $f = [x_1 \text{ on } A_1; \ldots; x_m \text{ on } A_m]$ and $g = [y_1 \text{ on } B_1; \ldots; y_n \text{ on } B_n]$,

$$f \succ g \Leftrightarrow U\left(\sum_{i=1}^{m} \pi(A_i)\, \delta_{x_i}\right) > U\left(\sum_{j=1}^{n} \pi(B_j)\, \delta_{y_j}\right).$$

We follow Machina and Schmeidler's (1992) rendering of Savage's theorem. The following set of axioms is necessary and sufficient for a preference relation to admit a Subjective Expected Utility representation.[17]

Axiom P1 (Ordering): *The relation $\succsim$ is complete, reflexive and transitive.*

Axiom P2 (Sure-Thing Principle) *For all events E and acts f, f^*, g and h,*

$$[f^* \text{ on } E;\ g \text{ on } E^c] \succsim [f \text{ on } E;\ g \text{ on } E^c]$$

$$\Rightarrow \quad [f^* \text{ on } E;\ h \text{ on } E^c] \succsim [f \text{ on } E;\ h \text{ on } E^c]$$

Axiom P3 (Eventwise Monotonicity) *For all outcomes x and y in $\mathcal{X}$, non-null events E and*

[17] Savage also has a state-wise dominance axiom (P7) but it is not needed in this setting in which every act has finite range.

acts g,

$$[x \text{ on } E;\ g \text{ on } E^c] \succsim [y \text{ on } E;\ g \text{ on } E^c] \Leftrightarrow x \succsim y.$$

Axiom P4 (Weak Comparative Probability) *For all events* A *and* B,*and outcomes* $x^* \succ x$ *and* $y^* \succ y$,

$$[x^* \text{ on } A;\ x \text{ on } A^c] \succsim [x^* \text{ on } B;\ x \text{ on } B^c]$$

$$\Rightarrow [y^* \text{ on } A;\ y \text{ on } A^c] \succsim [y^* \text{ on } B;\ y \text{ on } B^c]\,.$$

Axiom P5 (Nondegeneracy) *There exist outcomes* x *and* y *such that* $x \succ y$.

Axiom P6 (Small Event Continuity) *For any acts* $f \succ g$ *and outcomes* x, *there exists a finite set of events* $\{A_1, \ldots, A_n\}$ *forming a partition of* S *such that*

$$f \succ [x \text{ on } A_i;\ g \text{ on } A_i^c] \text{ and } [x \text{ on } A_i;\ f \text{ on } A_i^c] \succ g, \text{ for all } i \in \{1, \ldots, n\}\,.$$

The first and last axioms are analogous to the standard ordering and continuity axioms in choice under certainty. The fifth axiom simply requires the individual not be indifferent between all outcomes. The fourth axiom allows us to deduce from the agent's preferences over simple bets an induced relative likelihood relation over events. We shall see below that the third axiom is necessary for the induced risk preferences to satisfy first-order stochastic dominance.

As Machina and Schmeidler observe, **P2** may be viewed as being "expected utility-based". That is, not only does it play a critical role in ensuring that beliefs can be represented by a unique subjective probability measure but it also entails that the induced risk preferences are represented by a von Neumann-Morgenstern affine utility function. The aim of Machina and Schmeidler (1992) was to investigate whether it was possible to have the former and without necessarily entailing the latter.

Suppose we fix a probability measure π over events, then we can associate with any act $f = [x_1$ on $A_1; \ldots; x_m$ on $A_n]$, the lottery $\mathbf{R}_f^{\pi} = \sum_{i=1}^{m} \pi(A_i)\, \delta_{x_i}$, that is, the lottery formed by taking the $\pi-$weighted average of the degenerate lotteries corresponding to the outcomes in the range of f.

Probabilistic sophistication means that there exists a unique subjective probability measure π defined over the set of events, such that the individual ranks any pair of acts *only* on the basis of the lotteries they induce through π. Formally:

Probabilistic Sophistication A preference relation $\succsim$ on $\mathcal{F}$ is said to be *probabilistically sophisticated* if there exists a *unique* subjective probability measure $\pi(\cdot)$ over events such that for all pairs of acts f and g, if $\mathbf{R}_f^\pi = \mathbf{R}_g^\pi$ then $f \sim g$.

If an agent is probabilistically sophisticated, we can construct her induced *risk preferences* as follows: for any pair of lotteries $\mathbf{R}$ and $\mathbf{R}^*$, we set $\mathbf{R} \succsim_L \mathbf{R}^*$ if there exists a pair of acts $f \succeq g$ such that $\mathbf{R} = \mathbf{R}_f^\pi$ and $\mathbf{R}^* = \mathbf{R}_g^\pi$. Furthermore, if an agent is probabilistically sophisticated then knowledge of an individual's subjective beliefs π and her risk preferences $\succsim_L$ enables us to recover her entire preference relation over acts.

Assume that the agent's induced preferences over lotteries (her risk preferences) can be represented by a (not necessary affine) function $V(.)$. Then probabilistic sophistication (with respect to π) is equivalent to:[18] for any pair of acts f and g,

$$f \succ g \Leftrightarrow V\left(\mathbf{R}_f^\pi\right) > V\left(\mathbf{R}_g^\pi\right).$$

If we further assume that the agent's preferences satisfy **P2** then her risk preferences would satisfy independence and hence $\succsim_L$ (respectively, $\succsim$) would admit a (Subjective) Expected Utility representation. Hence in their axiomatization of subjective probability, Machina and Schmeidler (1992) drop **P2** but they need to strengthen **P4**.

Axiom P4* (Strong Comparative Probability) *For all pairs of disjoint events A and B, outcomes $x^* \succ x$ and $y^* \succ y$, and acts g and h,*

[18] In fact, Machina and Schmeidler (1992) state their definition of probabilistic sophistication in terms of the representation $V(\cdot)$. We favor working with a definition stated purely in terms of preference as the concept of probabilistic sophistication is logically independent of whether or not the preferences admit a functional representation.

$$[x^* \text{ on } A;\ x \text{ on } B;\ g \text{ on } (A \cup B)^c] \succsim [x \text{ on } A;\ x^* \text{ on } B;\ g \text{ on } (A \cup B)^c]$$

$$\Rightarrow \quad [y^* \text{ on } A;\ y \text{ on } B;\ h \text{ on } (A \cup B)^c] \succsim [y \text{ on } A;\ y^* \text{ on } B;\ h \text{ on } (A \cup B)^c].$$

Any probabilistically sophisticated agent whose preferences satisfy stochastic monotonicity must satisfy this axiom. For such an agent, the first preference indicates $\pi(A) \geq \pi(B)$, which in turn implies the second preference.

Like **P4**, **P4*** stipulates that the relative likelihood assessments of the disjoint events A and B are invariant to the prizes directly used in the bet. But Machina and Schmeidler's axiom imposes an additional invariance: these relative likelihood assessments are invariant to the outcomes that obtain on the states outside of A and B. In Savage, this invariance is derived from **P2** (The Sure-Thing Principle). This implies that weakening **P2** in a suitable way should be sufficient to derive the second invariance. We do this in the next section by weakening **P2** to an appropriate concept of Stochastic Monotonicity for acts.

3.3 Probabilities via Stochastic Monotonicity

We first consider two implications of Stochastic Monotonicity for risk preferences, and then extend these ideas to preferences over acts. Let $\succsim_L$ be a risk preference relation on the lotteries over $\mathcal{X}$. Stochastic Monotonicity for $\succsim_L$ may be expressed in terms of basic substitution operations on lotteries. One basic substitution is to take an outcome in the support of a lottery and change it to a worse outcome. Stochastic Monotonicity says we should prefer the lottery before this change.

Stochastic Monotonicity (I) For all $x, y \in \mathcal{X}$, all $\alpha \in (0, 1]$, and all $\mathbf{R} \in \mathcal{L}$

$$x \succ y \ \Leftrightarrow\ \alpha\delta_x + (1-\alpha)\,\mathbf{R} \succ \alpha\delta_y + (1-\alpha)\,\mathbf{R}. \tag{1}$$

Another basic substitution is to move a probability mass from a better to a worse outcome within the support of a roulette lottery. Again Stochastic Monotonicity says we should prefer the lottery before this change. Thus, if any two roulette lotteries differ only in the weight they assign to just

two outcomes, the agent should prefer the lottery that puts the larger weight on the better outcome.

Stochastic Monotonicity (II) For any pair of outcomes $x, y \in \mathcal{X}$, if two roulette lotteries $\mathbf{R}, \mathbf{R}^* \in \mathcal{L}$ have the property that $\mathbf{R}(z) = \mathbf{R}^*(z)$ for all $z \notin \{x, y\}$, then

$$\begin{aligned} & \alpha\mathbf{R} + (1-\alpha)\,\mathbf{R}^{**} \succsim \alpha\mathbf{R}^* + (1-\alpha)\,\mathbf{R}^{**} \\ \Rightarrow \quad & \alpha\mathbf{R} + (1-\alpha)\,\hat{\mathbf{R}}^{**} \succsim \alpha\mathbf{R}^* + (1-\alpha)\,\hat{\mathbf{R}}^{**} \end{aligned} \tag{2}$$

for all $\mathbf{R}^{**}, \hat{\mathbf{R}}^{**} \in \mathcal{L}$ and $\alpha \in (0, 1]$.

Notice that, in expression (2), $\mathbf{R}$ and $\mathbf{R}^*$ differ only in the probability they assign to x and y Similarly $\alpha\mathbf{R} + (1-\alpha)\,\mathbf{R}^{**}$ and $\alpha\mathbf{R}^* + (1-\alpha)\,\mathbf{R}^{**}$ (and $\alpha\mathbf{R} + (1-\alpha)\,\hat{\mathbf{R}}^{**}$ and $\alpha\mathbf{R}^* + (1-\alpha)\,\hat{\mathbf{R}}^{**}$) also differ only in the probability they assign to x and y. Furthermore $\mathbf{R}$ assigns more weight than $\mathbf{R}^*$ to the better outcome if and only if $\alpha\mathbf{R} + (1-\alpha)\,\mathbf{R}^{**}$ (respectively, $\alpha\mathbf{R} + (1-\alpha)\,\hat{\mathbf{R}}^{**}$) assigns more weight than $\alpha\mathbf{R}^* + (1-\alpha)\,\mathbf{R}^{**}$ (respectively, $\alpha\mathbf{R}^* + (1-\alpha)\,\hat{\mathbf{R}}^{**}$) to the better outcome.

Both these Stochastic Monotonicity axioms resemble Independence in that they preserve preference ordering under the substitution of (sub-)lotteries. But these axioms only involve lotteries that differ just on two outcomes. This is a considerable weakening in that all risk preferences that satisfy Stochastic Monotonicity satisfy both these axioms, regardless of whether or not they satisfy Independence.

The two Stochastic Monotonicity axioms are equivalent if the risk preference $\succeq_{\mathcal{L}}$ is continuous. However, their analogs for acts are not equivalent, so we will impose them separately.

Savage's axiom **P3** (Eventwise Monotonicity) is an analog of Stochastic Monotonicity (I) for acts. To see the analogy, let $g = [z \text{ on } E,\ z_1 \text{ on } E_1; \ldots; z_n \text{ on } E_n]$. Suppose the agent's probability assessments are given by π. Then the terms $[x \text{ on } E;\ g \text{ on } E^c]$ and $[y \text{ on } E;\ g \text{ on } E^c]$ in the statement of the axiom reduce to the lotteries $\pi(E)\,[\delta_x] + \sum_{i=1}^{n} \pi(E_i)\,\delta_{z_i}$ and $\pi(E)\,[\delta_y] + \sum_{i=1}^{n} \pi(E_i)\,\delta_{z_i}$, similar to the terms on the right side of expression (1). Therefore, if a Probabilistically Sophisticated agent satisfies Stochastic Monotonicity (I), she must satisfy **P3** (Eventwise Monotonicity).

The following is an analog of Stochastic Monotonicity (II) for acts. It restricts the Sure-Thing Principle to apply only if the two acts in question have the property that, for each state, the outcomes that the two acts assign can only differ when the acts are assigning one of two particular outcomes.

Axiom P2 (Two-Outcome Sure-Thing Principle)** *For any pair of outcomes $x, y \in \mathcal{X}$, any event E, and any acts f, f^*, g and h, if for all $z \notin \{x, y\}$, $f(s) = z$ if and only if $f^*(s) = z$ for all states $s \in S$, then*

$$[f \text{ on } E;\ g \text{ on } E^c] \succsim [f^* \text{ on } E;\ g \text{ on } E^c] \ \Rightarrow\ [f \text{ on } E;\ h \text{ on } E^c] \succsim [f^* \text{ on } E;\ h \text{ on } E^c]. \quad (3)$$

To see the analogy to Stochastic Monotonicity (II), suppose the agent's probability assessments are given by π. The implication trivially follows if E is null or its complement is null, so suppose $\pi(E) \in (0,1)$ and suppose $f = [z_1 \text{ on } A_1; \ldots; z_m \text{ on } A_m]$, $f^* = [z_1^* \text{ on } B_1; \ldots; z_n^* \text{ on } B_n]$, $g = [\hat{x}_1 \text{ on } C_1; \ldots; \hat{x}_{\hat{m}} \text{ on } C_{\hat{m}}]$ and $h = [\hat{y}_1 \text{ on } D_1; \ldots; \hat{y}_{\hat{n}} \text{ on } D_{\hat{n}}]$. Set

$$\mathbf{R} = \tfrac{1}{\pi(E)} \left[\textstyle\sum_{i=1}^{m} \pi(A_i \cap E)\, \delta_{z_i}\right], \qquad \mathbf{R}^{**} = \tfrac{1}{1-\pi(E)} \textstyle\sum_{k=1}^{\hat{m}} \pi(C_k \cap E^c)\, \delta_{\hat{x}_k},$$

$$\mathbf{R}^{*} = \tfrac{1}{\pi(E)} \left[\textstyle\sum_{j=1}^{n} \pi(B_j \cap E)\, \delta_{z_j^*}\right], \qquad \hat{\mathbf{R}}^{**} = \tfrac{1}{1-\pi(E)} \textstyle\sum_{\hat{k}=1}^{\hat{n}} \pi\left(D_{\hat{k}} \cap E^c\right) \delta_{\hat{y}_{\hat{k}}}.$$

Hence the reduction of $[f \text{ on } E;\ g \text{ on } E^c]$ is $\pi(E)\,\mathbf{R} + [1 - \pi(E)]\,\mathbf{R}^{**}$ and the reduction of $[f^* \text{ on } E;\ g \text{ on } E^c]$ is $\pi(E)\,\mathbf{R}^{*} + [1 - \pi(E)]\,\mathbf{R}^{**}$. Similarly, the reduction of $[f \text{ on } E;\ h \text{ on } E^c]$ is $\pi(E)\,\mathbf{R} + [1 - \pi(E)]\,\hat{\mathbf{R}}^{**}$ and the reduction of $[f^* \text{ on } E;\ h \text{ on } E^c]$ is $\pi(E)\,\mathbf{R}^{*} + [1 - \pi(E)]\,\hat{\mathbf{R}}^{**}$. Since, in each state, $f(s)$ and $f^*(s)$ may differ only if they assign the outcomes x or y, the 'conditional' reductions $\mathbf{R}$ and $\mathbf{R}^*$ differ only in the probabilities they assign to x and y. Therefore, if a Probabilistically Sophisticated agent satisfies Stochastic Monotonicity (II), she must satisfy the Two-Outcome Sure-Thing Principle..

Main result.

Theorem 4 The following two statements are equivalent:

(a) The preference relation $\succsim$ on $\mathcal{F}$ satisfies *Ordering*, *Two-Outcome Sure-Thing Principle*, *Eventwise Monotonicity*, *Weak Comparative Probability*, *Nondegeneracy* and *Small Event Continuity*.

(b) There exists a unique finitely-additive convex-ranged probability measure π defined over subsets of $\mathcal{S}$ and a non-constant, Mixture Continuous, Stochastically Monotonic function V on $\mathcal{L}$, such that for all pairs of acts f and g in $\mathcal{F}$

$$f \succsim g \Leftrightarrow V\left(\mathbf{R}_f^{\pi}\right) \geq V\left(\mathbf{R}_g^{\pi}\right).$$

Proof: It is enough to show that our axioms imply Machina & Schmeidler's axioms. In particular, we show that the conjuction of **P2**** and **P4** implies **P4***. To this purpose, consider two disjoint events A and B, outcomes $x^* \succ x$ and $y^* \succ y$, and act g such that

$$[x^* \text{ on } A;\ x \text{ on } B;\ g \text{ on } (A \cup B)^c] \succsim [x \text{ on } A;\ x^* \text{ on } B;\ g \text{ on } (A \cup B)^c] \tag{4}$$

We would like to show that for any act h,

$$[y^* \text{ on } A;\ y \text{ on } B;\ h \text{ on } (A \cup B)^c] \succsim [y \text{ on } A;\ y^* \text{ on } B;\ h \text{ on } (A \cup B)^c] \tag{5}$$

Let $f = [x^* \text{ on } A;\ x \text{ on } A^c]$ and $f' = [x \text{ on } A;\ x^* \text{ on } A^c]$, so that expression 4 implies $f_{(A \cup B)} g \succsim f'_{(A \cup B)} g$. Note that f and f' have only two outcomes in their range, so we can apply **P2**** to get $f_{(A \cup B)} x \succsim f'_{(A \cup B)} x$. Since $f_{(A \cup B)} x = [x^* \text{ on } A;\ x \text{ on } A^c]$ and $f'_{(A \cup B)} x = [x^* \text{ on } B;\ x \text{ on } B^c]$, **P4** implies $[y^* \text{ on } A;\ y \text{ on } A^c] \succsim [y^* \text{ on } B;\ y \text{ on } B^c]$. This is equivalent to $\bar{f}_{(A \cup B)} y \succsim \tilde{f}_{(A \cup B)} y$ where $\bar{f} = [y^* \text{ on } A;\ y \text{ on } A^c]$ and $\tilde{f} = [y \text{ on } A;\ y^* \text{ on } A^c]$. Note that $\bar{f}$ and $\tilde{f}$ have only two outcomes in their range. Now invoking **P2**** yields $\bar{f}_{(A \cup B)} h \succsim \tilde{f}_{(A \cup B)} h$, which is equivalent to 5, the desired expression.

References

[1] Aczel, J., 1987. A Short Course on Functional Equations. D. Reidel Publishing Company, Dordrecht

[2] Allais, M., 1953. Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Amèricaine. Econometrica 21: 503-546.

[3] Anscombe, F., Aumann, R., 1963. A Definition of Subjective Probability. Annals of Mathematical Statistics 34: 199-205.

[4] Banker, R., 1981. Equity Consideration in Traditional Full-Cost Allocation Pratices: An Axiomatic Perspective. In: Moriarty S (ed) Joint Cost Allocations. Norman, University of Oklahoma.

[5] Bergantiños, G., Vidal-Puga, J. J., 2004. A Fair Rule in Minimum Most Dspanning Tree Problems, forthcoming in Journal of Economic Theory.

[6] Bird, C. J., 1976. On cost allocation for a spanning tree: a game theoretic approach. Networks 6, 335-350.

[7] Brânzei, R., Moretti, S., Norde, H., Tijs, S., 2004. The P-value for Cost Sharing in Minimum Cost Spanning Tree Situations. Theory and Decision 56, 47-61.

[8] Chun, Y., 1988. The Proportional Solution for Rights Problems. Mathematical Social Sciences 15: 231-246.

[9] Chun, Y., 2004. Consistency and Monotonicity in Sequencing Problems. Mimeo.

[10] Chun, Y., 2006. A Pessimistic Approach to the Queuing Problem. Mathematical Social Sciences 51: 171-181.

[11] Claus, A., Kleitman, D. J., 1973. Cost allocation for a spanning tree. Networks 3, 289-304.

[12] Curiel, I., Pederzoli, G., Tijs, S., 1989. Sequencing Games. European Journal of Operational Research 40: 344-351

[13] Curiel, I., Hamers, H., Klijn, F., 2002. Sequencing Games: A Survey. In: Borm P, Peters H (eds) Chapters in Game Theory in Honor of Stef Tijs. Kluwer Academic Publishers, Boston, pp. 27-50

[14] De Frutos, M. A., 1999. Coalitional Manipulations in a Bankruptcy Problem. Review of Economic Design 4: 255-272.

[15] Dijkstra, E. W., 1959. A note on two problems in connection with graphs. Numerische Mathematik 1, 269-271.

[16] Douceur, J., 2002. The Sybil Attack. Proceedings of the 1st International Workshop on Peer-to-Peer Systems (IPTPS 2002), LNCS 2429: 251-260.

[17] Dutta, B., Kar, A., 2004. Cost monotonicity, consistency, and minimum cost spanning tree games. Games Econ. Behav. 48, 223-248.

[18] Feltkamp, T., Tijs, S., Muto, S., 1994. On the irreducible core and the equal remaining obligations rule of minimum cost spanning extension problems. Mimeo.

[19] Feltkamp, T., Tijs, S., Muto, S., 2000. Bird's tree allocations revisited. In: Patrone, F., Garcia-Jurado, I., Tijs, S. (Eds.), Game Practice: Contributions from Applied Game Theory. Kluwer, Dordrecht, pp. 75-89.

[20] Gómez-Rúa, M., Vidal-Puga, J.J., 2006. No advantegous merging in minimum cost spanning tree problems.

[21] Granot, D., Huberman, G., 1981. Minimum cost spanning tree games. Mathematical Programming 21, 1-18.

[22] Granot, D., Huberman, G., 1982. The relationship between convex games and minimum cost spanning tree games: a case for permutationally convex games. SIAM J. Alg. Disc. Meth. 3, 288-292.

[23] Granot, D., Huberman, G., 1984. On the core and nucleolus of minimum cost spanning tree games. Mathematical Programming 29, 323-347.

[24] Grant, S. Subjective Probability without Eventwise Monotonicity: or How Machina's Mom May also be Probabilistically Sophisticated. Econometrica 63, 159-189.

[25] Grant, S., Polak, B., 2006. Bayesian Beliefs with Stochastic Monotonicity: An Extension of Machina and Schmeidler, Journal of Economic Theory 130: 264-282.

[26] Hain, R., Mitra, M., 2004. Simple Sequencing Problems with Interdependent Costs. Games and Economic Behaviour 48(2): 271-291.

[27] Jarník, V. O., 1930. O Jistém problému minimálním. Práce Moravske Přírodovedëcke Spolecnosti 6, 57-63.

[28] Ju, B. G., 2003. Manipulations via Merging and Splitting in Claim Problems. Review of Economic Design 8: 205-21.

[29] Kar, A., 2002. Axiomatization of the Shapley value in minimum cost spanning tree games. Games Econ. Behav. 38, 265-277.

[30] Klijn, F., Sánchez, E., 2002. Sequencing Games without a Completely Specific Initial Order. Reports in Statistics and Operations Research, Report 02-04: 1-17.

[31] Machina, M., Schmeidler, D., 1992. A more robust definition of subjective probability, Econometrica 60, 745-780.

[32] Machina, M., Schmeidler, D., 1995. Bayes without Bernoulli. Journal of Economic Theory 67, 106-128.

[33] Maniquet, F., 2003. A Characterization of the Shapley Value in Queueing Problems. Journal of Economic Theory 109(1): 90-103.

[34] Megiddo, N., 1978. Cost allocation for Steiner trees. Networks 8, 1-6.

[35] Mitra, M., 2001. Mechanism Design in Queuing Problems. Economic Theory 17: 277-305

[36] Mitra, M., 2002. Achieving the First Best in Sequencing Problems. Review of Economic Design 7: 75-91

[37] Mitra, M., Sen, A., 1998. Dominant Strategy Implementation of First-Best Public Decisions. Indian Statistical Institute Delhi Centre, Manuscript No. 98-02.

[38] Moulin, H.,1985. Egalitarianism and Utilitarianism in Quasi-Linear Bargaining. Econometrica 53(1): 49-67.

[39] Moulin, H.,1987. Equal or Proportional Division of a Surplus, and Other Methods. International Journal of Game Theory 16(3): 161-186.

[40] Moulin, H., 2004a. On Scheduling Fees to Prevent Merging, Splitting and Transferring jobs. Forthcoming in Mathematics of Operations Research.

[41] Moulin, H., 2004b. Proportional Scheduling, Split-Proofness and Merge-Proofness. Forthcoming in Games and Economic Behavior.

[42] Norde, H., Moretti, S., Tijs, S., 2001. Minimum cost spanning tree games and population monotonic allocation schemes. European Journal of Operations Research 154(1), 84-97.

[43] Postlewaite, A., 1979. Manipulation via Endowments. The Review of Economic Studies 46: 255-262.

[44] Prim, R. C., 1957. Shortest connection network and some generalization. Bell System Technical Journal 36, 1389-1401.

[45] Queyranne, M., 1993. Structure of a Simple Scheduling Polyhedron. Mathematical Programming 58: 263-285.

[46] Savage, L. J., 1954. The Foundations of Statistics. Wiley, New York, 1954, Revised and enlarged, Dover Publications, New York, 1972.

[47] Sharkey, W. W., 1995. Network models in economics. In: Ball, M. O., Magnanti, T. L., Nonma, C. L., Nemhauser, G. L. (Eds.), Handbooks in Operation Research and Management Science, Elsevier, New York, pp. 713-765.

[48] Suijs, J., 1996. On Incentive Compatibility and Budget Balancedness in Public Decision Making. Economic Design 2: 193-209.

[49] Thomson, W., 1983. The Fair Division of a Fixed Supply Among a Growing Population. Mathematics of Operations Research 8: 319-326.

[50] Von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton: Princeton University Press, 60th Anniversary Edition, 2004.

Appendix: Proofs

Lemma 1

Take a w–quasi-proportional rule that satisfies scale invariance and demand monotonicity, and any scheduling problem (N, x). Scale invariance implies, for all $\gamma > 0$ and $i, j \in N$,

$$\begin{aligned} \theta(\gamma x_i, \gamma x_j) &= \theta(x_i, x_j) \\ &\Leftrightarrow \frac{w(\gamma x_i)}{w(\gamma x_i) + w(\gamma x_j)} = \frac{w(x_i)}{w(x_i) + w(x_j)} \\ &\Leftrightarrow \frac{w(\gamma x)}{w(x)} = \frac{w(\gamma \bar{x})}{w(\bar{x})} \text{ for all } x, \bar{x} > 0. \end{aligned}$$

Hence, there exists a function $f : \mathbb{R}_{++} \to \mathbb{R}_{++}$ such that

$$\frac{w(\gamma x)}{w(x)} = f(\gamma) \text{ for all } x > 0.$$

This implies

$$w(x^r) = f(x) \cdot w\left(x^{r-1}\right) \text{ for all } x > 0 \text{ and } r \in \mathbb{R}. \tag{1}$$

Setting $r = 1$ in this equation gives

$$w(x) = w(1) \cdot f(x) \text{ for all } x > 0. \tag{2}$$

Plugging in for $w(x^r)$ and $w\left(x^{r-1}\right)$ in (1) using (2), we get

$$f(x^r) = f(x) \cdot f\left(x^{r-1}\right) \tag{3}$$

Now, given any $z, a, b > 0, a \neq 1$, find $r_1, r_2 \in \mathbb{R}$ such that $a = z^{r_1}$ and $b = z^{r_2}$. Setting $x = z^{r_1}$ and $r = \frac{r_1 + r_2}{r_1}$ in (3), we get

$$f\left(\left(z^{r_1}\right)^{\frac{r_1+r_2}{r_1}}\right) = f\left(z^{r_1}\right) \cdot f\left(\left(z^{r_1}\right)^{\frac{r_1+r_2}{r_1}-1}\right)$$

which implies $f(a \cdot b) = f(a) \cdot f(b)$ for all $a, b > 0$. Now, defining $g(s) = \log f(e^s)$ we get

$$g(s+t) = g(s) + g(t) \text{ for all } s, t > 0$$

As our rule satisfies demand monotonicity, w is nondecreasing by Lemma 3 in Moulin [2004 b] and so is f because of (2). Hence, g is also nondecreasing, and by Corollary 1.8 in Aczel [1987], there exists a real constant α such that $g(s) = \alpha s$ for all $s > 0$. This in turn implies that $f(\gamma) = \gamma^\alpha$, and as f is nondecreasing, we have $\alpha \geq 0$. Equation (3) now becomes

$$w(x) = w(1) \cdot x^\alpha \text{ for all } x > 0.$$

Theorem 2

Take a w–quasi-proportional scheduling rule p^w such that w is concave and nondecreasing, and any scheduling problem (N, x). We assume without loss of generality that $S = \{1, 2\} \subset N$, and that $x_1 \leq x_2$. Suppose users 1 and 2 report t_1, t_2 such that $t_1 + t_2 = x_1 + x_2$. Since the ordering of the jobs (t_1, t_2) is irrelevant in relation to what is feasible for S, we restrict attention to transfers of the form $t_1 \leq t_2$.

An agreement specifies, for all orderings σ of N, the probability that user 1 will precede user 2 in the usage of the slots t_1, t_2. Hence, an agreement is a function $q : \Phi(N) \to \Delta(\Phi(S))$. We denote by y_i^q the expected wait time of user $i \in S$, induced by agreement q.

We proceed in two steps in order to show that p^w is pairwise transfer-proof. In step 1, we show that a sufficient condition for pairwise transfer-proofness is $x_1 y_1 + x_2 y_2 \leq x_1 y_1^0 + x_2 y_2^0$, where $y_i^0, i \in S$ denotes the expected wait of i, under the agreement that user 1 precedes user 2 in all the orderings. Note that y_1^0 and y_2^0 depend on the relative sizes of x_1 and t_1. If we have $t_1 < x_1$ and t_1 is before t_2 in the ex-post ordering of the jobs, for instance, x_1 will not fit in the slot t_1 and user 1 will have to wait until t_2 is served. In Step 2 we consider the three different cases regarding the relative sizes of x_1 and t_1, calculate the values of y_1^0 and y_2^0, and establish that the sufficient condition we derive in Step 1 holds.

Calculating the expected wait under agreement q is not trivial. One has to pay attention to the relative sizes of t_1 and x_1. The first step

Step 1 Without loss of generality, index the set $\Phi(N) = \{\sigma^1, \sigma^2, ..., \sigma^K\}$. We define q^0 to be the agreement under which 1 precedes 2 in all the orderings. Given an agreement q, we define a set of agreements $q^1, q^2, ..., q^K$ recursively, in the following manner:

For all $k = 1, 2, ..., K$:

$$q^k(\sigma) = q^{k-1}(\sigma) \text{ for all } \sigma \in \Phi(N) \setminus \sigma^k \text{ and } q^k(\sigma^k) = q(\sigma^k).$$

We see that $q^K = q$. For all $k = 1, 2, ..., K$, denote by y_i^k the expected wait of user $i, i \in S$, induced by agreement q^k.

Now suppose that $0 < t_1 < x_1$. If t_1 is before t_2 according to σ^k, moving from agreement q^{k-1} to q^k, the wait time of user 1 is increased by $(1 - q^k(\sigma^k))x_2$ and that of user 2 is decreased by $(1 - q^k(\sigma^k))x_1$ (recall that $x_1 \leq x_2$). So, if t_1 is before t_2 according to σ^k, y^k lies on the line that passes from y^{k-1} and has a slope of $-\frac{x_1}{x_2}$.As the absolute value of this slope is not greater than 1, y^k lies in the half space $H^k = \{(z_1, z_2) \in \mathbb{R}^2 : z_1 + z_2 \geq y_1^{k-1} + y_2^{k-1}\}$. If t_2 is before t_1 according to σ^k, moving from agreement q^{k-1} to q^k, the expected wait of user 1 is increased by $(1 - q_k(\sigma_k))\left(x_2 + \sum_{j \in M(\sigma^k)} x_j\right)$ where $M(\sigma)$ denotes the set of users whose jobs are served between t_1 and t_2 according to σ. From q^{k-1} to q^k the wait time of user 2 is decreased by $(1 - q^k(\sigma^k))\left(x_1 + \sum_{j \in M(\sigma^k)} x_j\right)$. Therefore, y^k lies on the line that passes from y^{k-1} and has a slope of $-\frac{x_1 + \sum_{j \in M(\sigma^k)} x_j}{x_2 + \sum_{j \in M(\sigma^k)} x_j}$, the absolute value of which is not greater than 1. This implies again that $y^k \in H^k$. Therefore, we have $y_1^0 + y_2^0 \leq y_1^K + y_2^K = y_1^q + y_2^q$, and a sufficient condition for pairwise transfer-proofness when $0 < t_1 < x_1$ is

$$y_1 + y_2 \leq y_1^0 + y_2^0 \tag{4}$$

Now suppose that $x_1 < t_1$. In this case, moving from agreement q^{k-1} to q^k, the expected delay of user 1 is increased by $(1 - q^k(\sigma^k))\left(x_2 + \sum_{j \in M(\sigma^k)} x_j\right)$, and that of user 2 is decreased by

$\left(1-q^k\left(\sigma^k\right)\right)x_1$ (recall that $t_1 \leq t_2$). So, y^k lies on the line that passes from y^{k-1} and has a slope of $-\frac{x_1}{x_2+\sum_{j\in M(\sigma^k)}x_j}$. As the absolute value of this slope is not greater than $\frac{x_1}{x_2}$, y^k lies in the half space

$$\tilde{H}^k=\left\{(z_1,z_2)\in\mathbb{R}^2 : x_1z_1+x_2z_2 \geq x_1y_1^{k-1}+x_2y_2^{k-1}\right\}$$

Therefore, we have $x_1y_1^0+x_2y_2^0 \leq x_1y_1^K+x_2y_2^K = x_1y_1^q+x_2y_2^q$ and a sufficient condition for pairwise transfer-proofness when $t_1 > x_1$ is

$$x_1y_1+x_2y_2 \leq x_1y_1^0+x_2y_2^0 \tag{5}$$

Step 2 We now show that p^w is pairwise transfer-proof. We have three cases to consider:

Case 1: $t_1 = 0$

This is the case where the two users "merge" their demand under one alias, and is dealt with in Moulin [2004b]. Proposition 2 in that paper states that a $w-$quasi-proportional rule is merge-proof if w is *nondecreasing.*

Case 2: $0 < t_1 < x_1$

We first calculate y^0, which is achieved when 1 precedes 2 in all orderings:

$$\begin{aligned} y_1^0 &= x_1+\textstyle\sum_{j\in N\setminus S}\theta\left(t_2,x_j\right)x_j \\ y_2^0 &= x_1+x_2+\textstyle\sum_{j\in N\setminus S}\left(1-\theta\left(x_j,\{t_1,t_2\}\right)\right)x_j \end{aligned}$$

where $\theta\left(x_j,\{t_1,t_2\}\right)=\frac{w(x_j)}{w(x_j)+w(t_1)+w(t_2)}$, $j\in N\setminus\{1,2\}$ is the probability that t_1 and t_2 are served before x_j.

Substituting for these values in inequality (4) gives the following sufficient condition for transfer-proofness:

$$\begin{aligned} &\sum_{j\in N\setminus S}\frac{w\left(x_1\right)}{w\left(x_1\right)+w\left(x_j\right)}x_j+\sum_{j\in N\setminus S}\frac{w\left(x_2\right)}{w\left(x_2\right)+w\left(x_j\right)}x_j \\ \leq\ &\sum_{j\in N\setminus S}\frac{w\left(t_2\right)}{w\left(t_2\right)+w\left(x_j\right)}x_j+\sum_{j\in N\setminus S}\frac{w\left(t_1\right)+w\left(t_2\right)}{w\left(t_1\right)+w\left(t_2\right)+w\left(x_j\right)}x_j \end{aligned}$$

As $\frac{w(x)}{w(x)+w(x_j)}$ is increasing in $w(x)$ and w is *nondecreasing*, $\frac{w(x_1)}{w(x_1)+w(x_j)} \leq \frac{w(t_2)}{w(t_2)+w(x_j)}$ for all x_j since $x_1 \leq t_2$ by assumption. Because w is *nondecreasing* we also have $w(x_2) \leq w(t_2) \leq w(t_1)+w(t_2)$, and this implies $\frac{w(x_2)}{w(x_2)+w(x_j)} \leq \frac{w(t_1)+w(t_2)}{w(t_1)+w(t_2)+w(x_j)}$ for all x_j. Hence, the above inequality holds.

Case 2: $t_1 > x_1$

In this case, y^0 is given as follows:

$$y_1^0 = x_1 + \sum_{j \in N \setminus S} \theta(\{t_1, t_2\}, x_j)\, x_j$$

where $\theta(\{t_1, t_2\}, x_j) = \frac{w(t_1)w(t_2)}{w(t_1)+w(t_2)+w(x_j)} \left(\frac{1}{w(t_1)+w(x_j)} + \frac{1}{w(t_2)+w(x_j)} \right)$ is the probability that x_j is served before t_1 and t_2.

$$y_2^0 = x_1 + x_2 + \sum_{j \in N \setminus S} (1 - \theta(x_j, \{t_1, t_2\}))\, x_j$$

Substituting these values in inequality (5) and simplifying we get the following sufficient condition for transfer-proofness

$$\begin{aligned} & x_1 \sum_{j \in N \setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} x_j + x_2 \sum_{j \in N \setminus S} \frac{w(x_2)}{w(x_2)+w(x_j)} x_j \\ \leq\ & x_1 \sum_{j \in N \setminus S} \frac{w(t_1)\, w(t_2)\, (w(t_1)+w(t_2)+2w(x_j))}{(w(t_1)+w(t_2)+w(x_j))\,(w(t_1)+w(x_j))\,(w(t_2)+w(x_j))} x_j \\ & + x_2 \sum_{j \in N \setminus S} \frac{w(t_1)+w(t_2)}{w(t_1)+w(t_2)+w(x_j)} x_j \end{aligned} \tag{6}$$

This inequality simply means that the externality from outside members does not increase after the transfers are reported. In the expression

$\frac{w(t_1)w(t_2)(w(t_1)+w(t_2)+2w(x_j))}{(w(t_1)+w(t_2)+w(x_j))(w(t_1)+w(x_j))(w(t_2)+w(x_j))}$, the fraction $\frac{w(t_1)+w(t_2)+2w(x_j)}{w(t_1)+w(t_2)+w(x_j)}$ is decreasing in $w(t_1)+w(t_2)$, which is maximized at $t_1 = t_2 = a = \frac{x_1+x_2}{2}$ because w is *concave*. So, this expression will not be smaller than $\frac{2w(a)+2w(x_j)}{2w(a)+w(x_j)}$. On the other hand, the expression $\frac{w(t_1)}{w(t_1)+w(x_j)}$ is increasing in $w(t_1)$ and hence also in t_1 as w is *nondecreasing*, so it is not smaller than $\frac{w(x_1)}{w(x_1)+w(x_j)}$. Now remem-

bering that $t_2 \geq \frac{x_1+x_2}{2}$, we similarly see that $\frac{w(t_2)}{w(t_2)+w(x_j)}$ is not smaller than $\frac{w(a)}{w(a)+w(x_j)}$. Note that $\frac{w(t_1)+w(t_2)}{w(t_1)+w(t_2)+w(x_j)}$ is increasing in $w(t_1)+w(t_2)$, so it is not smaller than $\frac{w(x_1)+w(x_2)}{w(x_1)+w(x_2)+w(x_j)}$ as w is *concave*. The above inequality will hold if it held when we replaced all these expressions by their minimized values. Substituting for these values we see that inequality (6) holds if

$$x_1 \sum_{j\in N\setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} x_j + x_2 \sum_{j\in N\setminus S} \frac{w(x_2)}{w(x_2)+w(x_j)} x_j$$
$$\leq x_1 \sum_{j\in N\setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} \frac{2w(a)}{2w(a)+w(x_j)} x_j + x_2 \sum_{j\in N\setminus S} \frac{w(x_1)+w(x_2)}{w(x_1)+w(x_2)+w(x_j)} x_j$$

Upon rearranging, this simplifies to

$$\sum_{j\in N\setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} x_1 x_j$$
$$\leq \sum_{j\in N\setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} \frac{2w(a)}{2w(a)+w(x_j)} x_1 x_j$$
$$+ \sum_{j\in N\setminus S} \frac{w(x_1)\cdot w(x_j)}{(w(x_1)+w(x_2)+w(x_j))} \frac{x_2}{w(x_2)+w(x_j)} x_j$$

Since w is *concave*, we have $w(x_2) - w'(x_2)\, x_2 \geq 0$. This in turn implies that $d\left(\frac{x_2}{w(x_2)+w(x_j)}\right)/dx_2 = \frac{w(x_2)+w(x_j)-w'(x_2)x_2}{(w(x_2)+w(x_j))^2} \geq 0$. This implies that, replacing x_2 with x_1 in the expression $\frac{x_2}{w(x_2)+w(x_j)}$ gives us a new sufficient condition which implies the original inequality (5):

$$\sum_{j\in N\setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} x_j \leq \sum_{j\in N\setminus S} \frac{w(x_1)}{w(x_1)+w(x_j)} \left[\frac{2w(a)}{2w(a)+w(x_j)} + \frac{w(x_j)}{(w(x_1)+w(x_2)+w(x_j))}\right] x_j$$

Note that for all x_j, $\frac{2w(a)}{2w(a)+w(x_j)} + \frac{w(x_j)}{w(x_1)+w(x_2)+w(x_j)} \geq 1$ since $\frac{2w(a)}{2w(a)+w(x_j)} \geq \frac{w(x_1)+w(x_2)}{w(x_1)+w(x_2)+w(x_j)}$ when w is *concave*. This implies that the previous inequality holds and establishes that p^w is pairwise transfer-proof.

Proposition 1

We show the only if part. Take any scheduling rule $p^\alpha, \alpha \geq 0$. We give an example of a scheduling problem to show that p^α is not transfer-proof for any $\alpha > 0$.

Consider the problem $N = (1,2,3,4), x = (1,1,4,2)$ and take $S = \{1,2,3\}$. Suppose the users in S report the transfers $t = (2,2,2)$. Denote by $y = (y_i)_{i\in N}$ the expected wait times in the problem (N,x), and let $2^\alpha = \kappa$. We have

$$y_1 = y_2 = 1 + \tfrac{1}{2}\cdot 1 + \tfrac{1}{\kappa^2+1}\cdot 4 + \tfrac{1}{\kappa+1}\cdot 2 = \tfrac{3}{2} + \tfrac{2\kappa^2+4\kappa+6}{\kappa^3+\kappa^2+\kappa+1}$$

$y_3 = 4 + 2 \cdot \frac{\kappa^2}{\kappa^2+1} \cdot 1 + \frac{\kappa^2}{\kappa^2+\kappa} \cdot 2 = 4 + \frac{4\kappa^3+2\kappa^2+2\kappa}{\kappa^3+\kappa^2+\kappa+1}$

Note that $\sum_{i=1}^{3} x_i y_i = 27 + \frac{8\kappa+4}{\kappa+1}$.

Now, under their untruthful report t, consider the following agreements among users 1,2 and 3:

Agreement 1 ($q1$):

$\{1,2\}$ precedes 3 in all the orderings of the jobs.

In each of the orderings 1 precedes 2 with probability 1/2.

Agreement 2 ($q2$):

3 precedes $\{1,2\}$ if x_4 is served third.

$\{1,2\}$ precedes 3 otherwise.

In each of the orderings 1 precedes 2 with probability 1/2.

Agreement 3 ($q3$):

3 precedes $\{1,2\}$ if x_4 is served third or last.

$\{1,2\}$ precedes 3 otherwise.

In each of the orderings 1 precedes 2 with probability 1/2.

When the users in S report the transfers $t = (2,2,2)$, the problem becomes (N, x_*) where $x_* = (2,2,2,2)$. As ρ^α is *anonymous*, all the orderings of N must have the same probability 1/4!. The following table summarizes the wait times of users 1, 2, and 3 under the above agreements:

	$q1$	$q2$	$q3$
$y_1 = y_2$	2	7/2	9/2
y_3	15/2	13/2	6

Denote by $y^{qk} = \left(y_1^{qk}, y_2^{qk}, y_3^{qk}\right)$ the wait times of users 1, 2, and 3 induced by agreement $qk, k = 1,2,3$. Note that there is a one-to-one relationship between α and y_3, and let $\alpha(y_3)$ denote the value of α that yields the expected wait y_3. It suffices to show that, for all $y_3 > 0$, there is a beneficial transfer maneuver. We consider two cases:

Case 1: α is such that $y_3 \leq y_3^{q1}$.

In what follows, we focus on user 3 and given different values of y_3, we construct a specific agreement among S such that, the expected wait of user 3 is precisely his expected wait in the

original problem. This agreement will be a convex combination of $q1, q2$, and $q3$. Once this agreement is constructed, straightforward comparisons yield that users 1 and 2 are better off, implying that $p^{\alpha(y_3)}$ is not transfer-proof. Note that $y_3^{q3} = 6 < y_3$ for all $\alpha > 0$, so that we consider two cases :

Case a: α is such that $y_3^{q3} < y_3 \leq y_3^{q2}$.

Define

$$y^{q(\beta)} = \beta y^{q2} + (1-\beta)\, y^{q3}$$

$y^{q(\beta)}$ is continuous in β, and we have $y_3^{q(1)} \geq y_3$ and $y_3^{q(0)} < y_3$. Therefore, by the intermediate value theorem, $\exists \beta^* \in (0,1] : y_3^{q(\beta^*)} = y_3$.

Note that $\sum_{i=1}^{3} x_i y_i^{q2} = \sum_{i=1}^{3} x_i y_i^{q3} = 33$, so that

$$\sum_{i=1}^{3} x_i y_i^{q(\beta^*)} = \beta^* \sum_{i=1}^{3} x_i y_i^{q2} + (1-\beta^*) \sum_{i=1}^{3} x_i y_i^{q3} = 33$$

Substituting for $y_3^{q(\beta^*)} = y_3 = 4 + \frac{4\kappa^3+2\kappa^2+2\kappa}{\kappa^3+\kappa^2+\kappa+1}$ in the above equation and using the fact that $y_1^{q(\beta^*)} = y_2^{q(\beta^*)}$, we find

$$y_1^{q(\beta^*)} = y_2^{q(\beta^*)} = \frac{17}{2} - \frac{8\kappa^3+4\kappa^2+4\kappa}{\kappa^3+\kappa^2+\kappa+1}$$

We now see that $y_1^{q(\beta^*)} = y_2^{q(\beta^*)} < y_1 = y_2$ holds iff

$$\begin{aligned} \frac{17}{2} - \frac{8\kappa^3+4\kappa^2+4\kappa}{\kappa^3+\kappa^2+\kappa+1} &< \frac{3}{2} + \frac{2\kappa^2+4\kappa+6}{\kappa^3+\kappa^2+\kappa+1} \\ &\Leftrightarrow \quad 0 < 2\kappa^3 - 2\kappa^2 + 2\kappa - 2 \end{aligned}$$

This inequality holds for all $\kappa > 1$, i.e. $\alpha > 0$. Hence, $y^{q(\beta^*)}$ is an improvement for users $1, 2$,and 3 over their original wait times y.

Case b: α is such that $y_3^{q2} < y_3 \leq y_3^{q1}$.

Define

$$y^{q(\delta)} = \delta y^{q1} + (1-\delta)\, y^{q2}$$

$y^{q(\delta)}$ is continuous in δ, $y_3^{q(1)} \geq y_3$ and $y_3^{q(0)} < y_3$. The intermediate value theorem implies that $\exists \delta^* \in (0,1] : y_3^{q(\delta^*)} = y_3$.

Note that $(x_1 + x_2 + x_4)\, y_1^{q1} + (x_3 + x_4)\, y_3^{q1} = 4 \cdot 2 + 6 \cdot \frac{15}{2} = 53$ and also $(x_1 + x_2 + x_4)\, y_1^{q2} + (x_3 + x_4)\, y_3^{q2} = 4 \cdot \frac{7}{2} + 6 \cdot \frac{13}{2} = 53$. This implies

$$(x_1 + x_2 + x_4)\, y_1^{q(\delta^*)} + (x_3 + x_4)\, y_3^{q(\delta^*)} = 53$$

Plugging in for $y_3^{q(\delta^*)} = y_3 = 4 + \frac{4\kappa^3+2\kappa^2+2\kappa}{\kappa^3+\kappa^2+\kappa+1}$, and using the fact that $y_1^{q(\delta^*)} = y_2^{q(\delta^*)}$, we find

$$y_1^{q(\delta^*)} = \frac{29}{4} - \frac{3}{2} \cdot \frac{4\kappa^3 + 2\kappa^2 + 2\kappa}{\kappa^3 + \kappa^2 + \kappa + 1}$$

We now see that $y_1^{q(\delta^*)} = y_2^{q(\delta^*)} < y_2 = y_1$ iff

$$\begin{aligned} \frac{29}{4} - \frac{3}{2} \cdot \frac{4\kappa^3 + 2\kappa^2 + 2\kappa}{\kappa^3 + \kappa^2 + \kappa + 1} &< \frac{3}{2} + \frac{2\kappa^2 + 4\kappa + 6}{\kappa^3 + \kappa^2 + \kappa + 1} \\ &\Leftrightarrow \quad 0 < \kappa^3 - 3\kappa^2 + 5\kappa + 1 \end{aligned}$$

This inequality holds for all $\kappa > 1$, and hence for all $\alpha > 0$. This implies that $y^{q(\delta^*)}$ is an improvement for users 1,2, and 3 over their initial wait times y.

Case 2: α is such that $y_3 > y_3^{q1}$.

The condition that $y_3 > y_3^{q1}$ implies that $\alpha > 1$. Now take any $w > 0$ such that $2\left(\frac{w+1}{2}\right)^{\alpha} < w^{\alpha}$. Consider the problem (N, x) with $N = \{1, 2, 3\}$ and $x = (1, w, z)$. Let y denote the wait profile in the problem (N, x). We have

$y_1 = 1 + \frac{1}{1+w^{\alpha}} w + \frac{1}{1+z^{\alpha}} z$

$y_2 = w + \frac{w^{\alpha}}{1+w^{\alpha}} 1 + \frac{w^{\alpha}}{z^{\alpha}+w^{\alpha}} z$

Take $S = \{1, 2\}$ and suppose S reports the transfers $(t, t), t = \frac{w+1}{2}$. Choose p such that

$$\frac{z^\alpha}{z^\alpha + 2t^\alpha} p = \frac{1}{1 + w^\alpha} \tag{7}$$

As $\alpha > 1$ and $2t^\alpha < w^\alpha$ by assumption, some $p < 1$ satisfies this equation for $z \geq 1$. Consider the agreement q for S under which 2 precedes 1 with probability p when 3 is last, and 1 precedes 2 otherwise. Denote by $y^q = (y_1^q, y_2^q)$ the expected delays for users 1 and 2 under this agreement. We compute

$$\begin{aligned} y_2^q &= x_2 + prob\{3 \text{ is not last}\}(x_1 + x_3) + prob\{3 \text{ is last}\}(1 - p) x_1 \\ &= 3 + \frac{2t^\alpha}{z^\alpha + 2t^\alpha}(1 + z) + \frac{z^\alpha}{z^\alpha + 2t^\alpha}(1 - p) 1 \\ &= 3 + \frac{2t^\alpha}{z^\alpha + 2t^\alpha} z + \frac{w^\alpha}{1 + w^\alpha} \end{aligned}$$

where we make use of (7) in the last step. Hence,

$$\begin{aligned} y_2^q &< y_2 \\ &\iff \frac{2t^\alpha}{z^\alpha + 2t^\alpha} < \frac{w^\alpha}{z^\alpha + w^\alpha} \\ &\iff 2t^\alpha < 3^\alpha \end{aligned}$$

which holds by assumption.We also have

$$\begin{aligned} y_1^q &= x_1 + prob\{3 \text{ is last}\} p x_2 + prob\{3 \text{ is first}\} x_3 \\ &= 1 + \frac{z^\alpha}{z^\alpha + 2t^\alpha} p3 + \frac{2t^\alpha}{z^\alpha + 2t^\alpha} \cdot \frac{t^\alpha}{z^\alpha + t^\alpha} z \end{aligned}$$

Making use of (7), we see $y_1^q < y_1$ holds iff

$$\frac{2t^\alpha}{z^\alpha + 2t^\alpha} \cdot \frac{t^\alpha}{z^\alpha + t^\alpha} < \frac{1}{1 + z^\alpha}$$

which holds for big z as $\alpha > 1$. Hence, the above agreement is beneficial for S when z is sufficiently large.