Split-proof probabilistic scheduling

Herve Moulin
Rice University

August 2004

Abstract

If shortest jobs are served first, splitting a long job into smaller jobs reported under different aliases, will reduce the actual wait until completion. If longest jobs are served first, the dual maneuver of merging several jobs under a single reported identity is profitable. Both manipulations can be avoided if the scheduling order is random, and users care only about the expected wait until completion of their job.

In the natural class of *separable* scheduling rules, Merge-proofness holds if individual delay is monotonic in own job size. Split-proofness is more demanding.

The Proportional rule stands out among rules immune to splitting and merging. It draws the job served last with probabilities proportional to sizes, then repeats among the remaining jobs. It is the only split-proof scheduling rule constructed in this *recursive* way, that minimizes the worst expected delay of individual jobs; or such that an agent with a longer job incurs a longer delay.

Key words: probabilistic scheduling, merging, splitting, proportional rule.

Acknowledgements: This work was prepared during successive visits to the Universities of Auckland, New Zealand, and Namur, Belgium, as well as the New School of Economics in Moscow. I am grateful to these institutions for their hospitality.

1 The problem and the main results

When processing jobs of different lengths, the familiar utilitarian goal of minimizing total waiting time requires to schedule them from shortest to longest. A user with a long job can lower his waiting time by splitting his order in several smaller jobs, if he can pretend that each small job is requested by a different person. The success of this maneuver rests on the unability to monitor the identity of the real beneficiaries of any job in the queue. This may reflect a feasibility constraint, such as the dispersion of information about the private benefits of each job and the prohibitive cost to retrieve it. Alternatively, the system designer may need to protect the users' privacy when jobs deal with sensitive information. Examples include users sharing a data base, which can be accessed on a single phone line, or a single decoding machine processing confidential documents. In large networks such as the Internet the proliferation of aliases is an issue of serious concern (Douceur [2002]). We study simple randomized scheduling disciplines invulnerable to the strategic faking of identities and the resulting tradeoffs with other normative concerns, such as the minimization of the worst case waiting time, and the fair distribution of the delay externality.

We focus on the two dual strategic maneuvers of splitting a single job in smaller jobs reported to the server under different aliases, and of merging several jobs into a single large job. In line with most of the scheduling literature (e.g., Lawler et al. [1993]), we assume that partially completed jobs are useless, so that an efficient server processes jobs whole[1]. Splitting is the cooperative manipulation where agent A who needs a job of size x, calls to help several agents A',A",.., not involved in the initial scheduling problem, and requests on their behalf jobs of sizes $a, a', a", ..$ such that $a+a'+a"+\cdots = x$. If the last agent served among A,A',A",.., is strictly earlier than A in the initial problem, the maneuver is profitable. Symmetrically agents A,B,C,.., who need jobs of sizes $a, b, c, ..$, can merge into a single agent A requesting a job of size $x = a + b + c + \cdots$. The merged agents schedule their true jobs as they please during the time where job x is processed. Merging is profitable if the waiting time of at least one merged agent can thus be reduced, without increasing that of any other.

A rule serving shortest jobs first is vulnerable to splitting, irrespective of

[1] At least in the scheduling problem, where all jobs are born at the same date. In the queuing problem, efficiency is compatible with preemptive service: see Section 8.

the tie-breaking rule; on the other hand it is "merge-proof", because a merged job is never served earlier than any of its component jobs. Similarly, serving longest jobs first (hence *maximizing* total waiting time) is a "split-proof" rule, whereas it is highly vulnerable to merging.

Call a scheduling rule *deterministic* if it selects a processing order for any profile of (deterministic) job sizes. A simple impossibility result holds true with as few as three agents: a deterministic rule cannot be both split-proof and merge-proof. Assume without loss of generality that when agents A,B,C each have a job of size 2, the rule orders them alphabetically. If now A,B show up with jobs $(a_1, b_1) = (2, 4)$, A must be served first, otherwise B and C both reduce their wait by merging in the initial three person problem. If B,C show up with $(b_2, c_2) = (4, 2)$, B must be served first otherwise B shortens his wait by splitting into $(a', b') = (2, 2)$. Next consider the problem $(a, b, c) = (1, 4, 1)$ If B is served last there, then at (b_2, c_2), C splits advantageously to $(a", b") = (1, 1)$. If B is served first or second at (a, b, c), and A,C merge into $a_1 = 2$, C benefits strictly whereas A suffers no harm (and even benefits if B was served first at (a, b, c)). This proves our claim.

Randomizing the scheduling of jobs is the simplest way to restore fairness when efficiency compels to process them whole. We submit another advantage of randomization: if we assume that each participant seeks to minimize the expected wait until the completion of his or her job,we can construct scheduling rules that are simultaneously split-proof and merge-proof. An example is the *Uniform* scheduling rule chooses each ordering of the n participants with equal probability $1/n!$, ignoring all differences in job lengths. It is easy to check its invulnerability to splitting and merging maneuvers[2]

Yet for all its simplicity, the Uniform rule is problematic on two accounts: *responsiveness* and *liability.*

Responsiveness is the idea –familiar in the cost sharing literature: e.g., Moulin and Sprumont [2003]– that waiting costs should be positively correlated with job lengths, because the longer the job, the larger the delay externality upon other participants. Loosely speaking, Shortest Jobs First is the most responsive rule of all; in particular it meets the two following responsiveness properties. The *Demand Monotonicity* axiom requires that the net waiting cost $y_i - x_i$, where y_i is the expected completion date of agent i's job with size x_i, be non decreasing in x_i. Longest Jobs First emphatically fails this property; the Uniform rule barely meets it, as $y_i - x_i = \frac{1}{2} \sum_{j \neq i} x_j$ does

[2] We omit the argument for brevity. The claim also follows from Proposition 4.

not depend on x_i at all. The *Ranking* axiom compares net waiting costs of any two agents i, j in a given problem: if $x_i \leq x_j$, it requires $y_i - x_i \leq y_j - x_j$. Although these two properties are not logically related, their normative content is similar: a larger job creates more delay for other users, hence it should bear a larger share of the delay externality. Ranking conveys this idea by interpersonal comparisons in a given problem, while Demand Monotonicity compares the shares of a given agent for different job sizes. Both Longest Jobs First and the Uniform rule fail Ranking. More generally in our model, Demand Monotonicity is "easy" to meet, and for many scheduling rules it implies Merge-proofness (Proposition 2); Ranking, on the other hand, is much more demanding and leaves little room for Split-proofness (see sections 6,7).

The liability of a scheduling rule measures the worst expected wait of a given job when other jobs are arbitrarily large. A job "chokes" if large jobs are scheduled ahead of it often enough that its expected wait grows unbounded. The risk of choking may discourage participation when users can opt it of the system (as discussed in Friedman et al. [2003]). To avoid choking, we place a finite cap on the expected wait of any job that only depends upon this job' size and the number of other users, but not on the size of these competing jobs. Under Shortest Jobs First, if n agents share the server the liability of a job of length x_i is $n \cdot x_i$: indeed the longest wait occurs when all other jobs are only slightly shorter than x_i. On the other hand, under either Longest Jobs First or the Uniform rule, this liability is infinite because y_i goes to infinity as x_j does, for any $j \neq i$.

We show that Merge-proofness is very easy to achieve in probabilistic scheduling, in particular it does not restrict the responsiveness of the rule, and is compatible with the minimal feasible liability[3]. Split-proofness, on the other hand, limits the responsiveness of the rule, and places a lower bound on its liability. Our main result is that a certain probabilistic scheduling rule achieves the best feasible compromise between split-proofness and the two goals of maximizing responsiveness and minimizing individual liability. This *Proportional* rule is split-proof and merge-proof, meets Monotonicity and Ranking, and guarantees the finite liability $n \cdot x_i$ at all problems. Moreover, within the set of *separable* rules or within that of *recursive* rules (see next paragraph), the Proportional one is characterized by the combination of Split-proofness, Merge-proofness, and either Ranking or the cap $n \cdot x_i$ on individual liability(Theorems 1 and 2 in Section 7).

[3]The smallest feasible liability is $\frac{n+1}{2} x_i$: Lemma 7 in Section 6.2.

A scheduling rule is separable if the relative ranking of any subset of jobs is independent of other jobs' sizes. A rule is recursive if it can be generated by n independent successive draws, choosing first the agent served *last*, then the agent served next to last in the reduced problem with $n-1$ agents, and so on.

The paramount example of separable rules are the *parametric* scheduling rules. Choose for each positive job length x, a cumulative distribution function F_x on $[0,+\infty[$. Given a set N of participants and a profile $x_i, i \in N$ of job sizes, we draw $|N|$ independent random variables Z_i, where Z_i has cdf F_{x_i}; then we order the jobs as the realizations of these variables, breaking ties with a uniform dice. For instance the Uniform rule obtains if the cdf F_x is independent of x. The Proportional rule is parametric with the Weibull cdf $F_x(z) = z^x$ for all $x > 0$. Its name comes from the fact that for any agents $i, j, k, ..$, the probability that i is served last equals $\frac{xi}{x_i+x_j+x_k+\cdot\cdot}$. More generally, a *quasi-proportional* rule is parametric with $F_x(z) = z^{w(x)}$ for all $x > 0$, where w is an arbitrary positive function. Such a rule is both separable and recursive; in fact the two properties Separability and Recursivity essentially characterize the family of quasi-proportional rules (Proposition 1).

A quasi-proportional rule is merge-proof if w is non decreasing; it is split-proof if and only if w is subadditive (Proposition 4). Ranking of a finite liability, on the other hand, essentially require w to be superadditive (Section 6): the pivotal role of the Proportional rule, for which w is additive, is thus especially clear in the quasi-proportional family.

Two general remarks about the interpretation of Merge-proofness and Split-proofness. If the context makes it easy to assume fake identities, splitting can be viewed as a non-cooperative manipulation, in contrast to merging that requires the cooperation of several different agents. In this case splitting is potentially a much more serious problem than merging. On the other hand, if the system manager can detect falke identities, a splitting maneuvers is again a cooperative move requiring the coordinated action of the agent with a "true" job with one or more "passive" agents who in truth have no job to submit. In that case splitting and merging are equally difficult to implement.

Next we note that even if the system manager cannot detect fake identities, agents must not be allowed to leave the system before the reported jobs are completed, otherwise any agent could "duplicate" his true job of length a in several requests for jobs of the same size (under assumed identities), and leave the system as soon as one of these jobs is completed. Oviously such maneuvers would always be profitable at those profiles x where all true jobs are of the same size (and at many other profiles as well).

In Section 3 we define scheduling *rules*, choosing a random ordering of service for every set of agents with jobs of arbitrary sizes, and scheduling *methods* computing the profile of expected waiting times on the same domain. The properties of Separability and Recursivity are introduced in Section 4, along with the important classes of parametric and quasi-proportional rules. In Section 5 we define merge-proof and split-proof rules, and show that these properties take a simple form for separable, and in particular parametric and quasi-proportional, rules. Section 6 defines the Ranking property, and the measure of individual liability. Our main results are in Section 7. Within the set of recursive rules, we characterize the Proportional one by the neutrality of splits (the property we call *Split-invariance*), or by the combination of Split-proofness and either the Ranking property or a simple bound on liability (Theorem 2). Parallel yet weaker results hold within the set of separable rules: under the same premises the scheduling *method* must be the Proportional one, but the rule may be different. The extension of some of our results to queuing, as well as some open questions are gathered in Section 8. The longer or peripheral proofs are in the Appendix.

2 Related literature

This paper is inspired by three essentially independent streams of microeconomic literature: the first one applies the mechanism design approach to queuing and scheduling, the second one discusses splitting and merging maneuvers in various fair division problems, and the third one studies the random assignment of private commodities.

We start with the research on on scheduling and queuing (under the assumption that only completed jobs matter) when monetary transfers are feasible. That literature typically assumes quasi-linear preferences in money, and linear waiting costs. One idea is to propose fair monetary compensations by applying solutions concepts from cooperative game theory: the Shapley value is increasingly popular for the scheduling model (Curiel et al. [1989],[1993],[2002], Klijn and Sanchez [2002], Maniquet [2003], Chun [2004 a,b]); in the scheduling problem, Haviv and Ritov [1998], Haviv [2001] apply the Aumann-Shapley pricing rule to various service disciplines. A different, and more developed, theme is to design cash transfers ensuring the truthful revelation of waiting costs. This leads to interesting (in particular, budget-balanced) Vicrey-Clarke-Groves mechanisms, or generalizations thereof: Dolan [1978], and more recently Suijs [1996], Mitra [2001,2002], Mitra and Sen [2002]. Closer to our model, Kittsteiner and Moldovanu[2003 a,b] consider the truthful revelation of job sizes, assuming that agents cannot be punished ex post for misreporting. This type of noncooperative misrepresentation differs sharply from the cooperative moves of splitting and merging jobs. In our model the system manager can punish agents who minimize ex ante the length of their job, and agents have no incentive to exagerate the size of their job either.

The companion paper Moulin [2004] discusses merging and splitting maneuvers in the quasi-linear scheduling model just discussed. The results there are less encouraging than in the probabilistic model of this paper, in the sense that no efficient mechanism can be both split-proof and merge-proof. Just like here, Merge-proofness proves much easier to meet than Split-proofness, that forces several unpalatable normative features, such as the systematic violation of Monotonicity and Ranking. See section 8 for further discussion of this model.

We turn to the discussion of splitting and merging in the fair division literature. The earliest contribution deal with the rationing problem (Banker [1981], Moulin [1987], De Frutos [1999], Ju [2003]) where the proportional rule is the only one immune against such maneuvers. In the quasi-linear social choice problem, the same property leads to the egalitarian division of surplus (Moulin [1985], Chun[]), and in the cost sharing problem with variable demands, to the Aumann-Shapley rule (Sprumont [2004]). Most of these results are surveyed in Ju and Myiagawa [2003].

Although splitting and merging are not meaningful there, the recent literature on random assignment is distinctly related to the present paper. Note that the scheduling problem is similar to the assignment of indivisible objects, where objects are slots in the processing order. The twist is that when jobs are of different sizes, the value of a given slot depends on who is served before, hence a slot is not a purely private good as in the assignment problem.The analog of our Uniform rule in the assignment model is the Random Priority mechanism, selecting a "pecking order" with uniform probability. Its efficiency and strategy-proofness properties are discussed by Abdulkadiroglu and Sonmez [1999], Bogomolnaia and Moulin [2001], who contrast it with the alternative Probabilistic Serial mechanism. The probabilistic scheduling of jobs of identical sizes is a special case of the assignment problem, for which the Probabilistic Serial unambiguously dominates Random Priority (Cres and Moulin [2000], Bogomolnaia and Moulin [2002]).

3 The model

The infinite set $\mathcal{N}$ of potential agents is fixed throughout. A *scheduling problem* is a pair (N, x) where N is a finite subset of $\mathcal{N}$, and x is a profile of non-negative job sizes, $x_i \in \mathbb{R}_+$ for all $i \in N$. The set of orderings of N is $\Phi(N)$, with generic element σ : $\sigma(i) < \sigma(j)$ means that job i is served/scheduled before job j. A random ordering is a probability distribution p on $\Phi(N)$; the set of such distributions is denoted $\Delta[\Phi(N)]$.

We assume that agents care only to minimize expected completion time of their own job, i.e., the expected wait y_i of agent i is her disutlity. Given the random ordering p and a problem (N, x), the expected wait of agent i until completion of her job is

$$y_i = x_i + \sum_{j \in N \diagdown i} prob\{\sigma(j) < \sigma(i)|x\} \cdot x_j \tag{1}$$

where $prob\{\sigma(j) < \sigma(i)|x\} = \sum_{\sigma:\sigma(j)<\sigma(i)} p_\sigma(N, x)$. Conversely, our first result describes those profiles $y \in \mathbb{R}^N_+$ *feasible* at a given problem (N, x), namely such that (1) holds for some lottery $p \in \Delta[\Phi(N)]$. Define for all $x \in \mathbb{R}^N_+$ and all $S \subseteq N$, the function $v(S, x) = \sum_S x_i^2 + \sum_{S(2)} x_i \cdot x_j$, where $S(2)$ is the set of non ordered pairs from S (the cardinality of this set is$\frac{|S|\cdot(|S|-1)}{2}$). Note that v is supermodular with respect to S.

Lemma 1

i) The profile $y \in \mathbb{R}^N_+$ is feasible at (N, x), $x \in \mathbb{R}^N_+$, if and only if for all i $\{x_i = 0 \Longrightarrow 0 \leq y_i \leq \sum_N x_j\}$, and moreover y belongs to the core of the game $(N, v(\cdot, x))$, i.e.,

$$\sum_N x_i \cdot y_i = v(N, x) \text{ and } \sum_S x_i \cdot y_i \geq v(S, x) \text{ for all } S \subseteq N.$$

ii) The profile $y \in \mathbb{R}^N_+$ is efficient at (N, x), $x \in \mathbb{R}^N_+$, if and only if for all i $\{x_i = 0 \Longrightarrow y_i = 0\}$, and moreover y belongs to the core of the game $(N, v(\cdot, x))$. We denote by $F(N, x)$ the set of efficient profiles y at (N, x).

This result is proven in Queyrane [1992]. For the sake of completeness we provide a (different) proof in the Appendix.

Lemma 1 implies that if all jobs are positive, *all* random orderings are efficient, because the weighted sum $\sum_N x_i \cdot y_i$ is independent of the choice of p. When some jobs are null efficiency only requires to schedule all null jobs before any non-null job.

We write $\mathbb{R}^N_{++}$ for the set of profiles x with all positive coordinates, and $x \gg 0$ as a short-hand for $x \in \mathbb{R}^N_{++}$. Similarly $x \geq 0$ is a short-hand for $x \in \mathbb{R}^N_+$.

Definition 1

An efficient *scheduling rule* is a mapping ρ_0 associating to each problem $(N,x), x \gg 0$, a random ordering $p \in \Delta[\Phi(N)]$.

An efficient *scheduling method* is a mapping μ associating to each problem $(N,x), x \geq 0$, an efficient profile $y \in F(N,x)$.

We only need to define efficient scheduling rules over strictly positive profiles x, because the relative ordering of null jobs is irrelevant, as long as they are served before all non null jobs. On the other hand, an efficient scheduling method is defined for all profiles, because the wait of null jobs is unambiguous; moreover it will prove convenient in Sections 5,6 to define a method for all profiles in $\mathbb{R}^N_+$.

In the sequel when we speak of a method or a rule, we always mean that it is efficient.

4 Separable, recursive and parametric rules

For a given set of n agents, the dimension of the set of scheduling rules is of the order of $n^{\frac{n^2}{2}}$, and in such a large set the properties of Merge- and Split-proofness are not enough to pin down some simple rules. We shall restrict attention to those rules meeting either Separability, a property reminiscent of the familiar Consistency axiom of the fair division literature, or a certain Recursivity property (Definition 6) below.

Given $N, S, S \subset N$, and $x \in \mathbb{R}^N_+$, we write $x_{[S]}$ its projection on $\mathbb{R}^S_+$; for any $\sigma \in \Phi(N)$, $\sigma[S] \in \Phi(S)$ is similarly the restriction of σ to S.

Definition 2

The scheduling rule ρ is *separable* if for all $N, S, S \subset N$, $x \gg 0$, the (random) ordering of the jobs in S is independent of the jobs outside S:

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

For a separable rule ρ, we speak of the probability that a job of size x_j for agent j precedes a job of size x_i for agent i, without specifying either the rest of the participants or the size of their jobs: we write this probability $\theta^{i,j}(x_i, x_j)$, so that the method μ associated with ρ takes the form

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

Definition 3

The scheduling method μ is *separable* if it takes the form (2) for all $N, x, x \geq 0$, where for all $i, j \in \mathcal{N}$, $\theta^{i,j}$ is a function from $\mathbb{R}^2_+$ into $[0,1]$ such that for all $a, b \geq 0$

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

It follows from Lemma 1 that for any choice of the functions $\theta^{i,j}$, equation (2) defines a feasible profile of expected waits: $\theta^{i,j}$ non negative implies $\sum_S x_i \cdot y_i = v(S,x) + \sum_{i \in S, j \in N \setminus S} \theta^{i,j}(x_i, x_j) \cdot x_i \cdot x_j \geq v(S,x)$. Yet it does not follows that we can always achieve the method $\{\theta^{i,j}\}$ by means of a *separable* scheduling rule ρ[4].

The family of parametric rules is an important class of separable scheduling rules that are *anonymous* as well, i.e., such that for all N, the mapping $x \to p(N,x)$ is symmetric in all variables on $\mathbb{R}^N_{++}$. In particular $\theta^{i,j} = \theta$ is independent of i, j. Choose for each $a > 0$ a cumulative distribution function F_a on $[0, +\infty[$ with no mass at 0. Thus F_a is any non decreasing and right-continuous function on $[0, +\infty[$ such that $F_a(0) = 0$ and $F_a(\infty) = 1$.

Definition 4

Given a scheduling problem $(N,x), x \gg 0$, the *parametric* rule associated with the family $\{F_a, a > 0\}$ picks $|N|$ independent random variables Z_{x_i} with cdf F_{x_i} for all $i \in N$, and orders jobs according to the realization of these variables, breaking ties randomly with uniform probability.

[4]For a given matrix $[\theta^{i,j}(x_i, x_j)] = [t_{i,j}]$ when can we find a lottery $p \in \Delta(\Phi(N))$ such that $prob\{\sigma(j) < \sigma(i)\} = t_{i,j}$ for all i, j ? Clearly the equalities $t_{i,j} + t_{j,i} = 1$ are necessary but not sufficient. To provide a complete set of necessary and sufficient conditions is still an open problem: see Fishburn [1992].

Note that if the problem (N, x) contains some null jobs, we can extend the definition of the parametric rule by drawing $Z_0 = 0$ for all null jobs, i.e., using the cdf F_0 concentrated at 0, $F_0(z) = 1$ for all $z \geq 0$.

The mutual stochastic independence of the random variables Z_{x_i} implies separability. Moreover the function $\theta^{i,j} = \theta$ is

$$\theta(a, b) = prob\{Z_a < Z_b\} + \frac{1}{2}prob\{Z_a = Z_b\} \text{ for all } a, b. \tag{3}$$

Examples of parametric rules include the three benchmark rules discussed in Section 1:

- Shortest Jobs First: F_a is the cdf concentrated at a ($F_a(z) = 1$ if $z \geq a$, $= 0$ if $z < a$); $\theta(a, b) = 1$ if $b < a$.
- Longest Jobs First: F_a is the cdf concentrated at $\frac{1}{a}$; $\theta(a, b) = 1$ if $a < b$.
- the Uniform rule: $F_a = F_1$ is any cdf with no mass at zero; $\theta(a, b) = \frac{1}{2}$ for all a, b.

Our next example is a rule where if Ann's job is larger than Bob's, the probability that Ann is scheduled after Bob is strictly above $\frac{1}{2}$ and strictly below 1.

Example 1. *The Serial rule.*
The cdf F_a is uniform on $[0, a]$, i.e., $F_a(z) = \min\{\frac{z}{a}, 1\}$. Here (3) gives easily $\theta(a, b) = \frac{a}{2b}$ if $a \leq b$, $= 1 - \frac{b}{2a}$ if $b \leq a$. Thus for a profile of job sizes such that $0 < x_1 \leq x_2 \leq \cdots \leq x_n$, the expected wait of agent i is

$$y_i = (1 + \frac{n-i}{2}) \cdot x_i + \sum_{j=1}^{i-1}(1 - \frac{x_j}{2x_i}) \cdot x_j$$

This formula explains the serial terminology. Indeed the expected wait of agent i does not depend on the sizes of jobs larger that his own job : if x and x' only differ in coordinate j, $j \neq i$, and $x_i \leq x_j < x'_j$, then $y_i(x) = y_i(x')$. This is the *serial principle* discussed by Sprumont [1998] and others. Moreover, the Serial scheduling method is the only mapping $x \to y$ from $\mathbb{R}_+^N$ into itself meeting the following four properties: the serial principle, equal treatment of equals ($x_i = x_j \Rightarrow y_i = y_j$), $\sum_N x_i \cdot y_i = v(N, x)$ (Lemma 1). We omit the straightforward proof for brevity.

We turn to the subfamily of parametric rules that is at the heart of this paper.

Definition 5
Let w be a function on $\mathbb{R}_+$, such that $w(0)=0$ and $a>0 \Longrightarrow w(a)>0$. The $w-$*quasi-proportional* scheduling rule is the parametric rule where for all $a>0$, F_a is the Weibull cdf $F_a(z)=\min\{z^{w(a)},1\}$ for all $z\geq 0$.

Lemma 2
Fix N, an ordering $\sigma \in \Theta(N)$, and a profile $x \in \mathbb{R}^N_{++}$. Set $w_k = w(x_{\sigma^{-1}(k)})$, so w_1 is the weight of the job scheduled first and w_n that of the job scheduled last. The probability of σ for the $w-$quasi-proportional rule is

$$p_\sigma = \frac{w_n}{\sum_1^n w_k} \cdot \frac{w_{n-1}}{\sum_1^{n-1} w_k} \cdot \ldots \cdot \frac{w_2}{w_1+w_2} \tag{4}$$

In particular $\theta(a,b)=\frac{w(a)}{w(b)+w(a)}$, and

$$y_i = x_i + \sum_{N\smallsetminus i} \frac{w(x_i)}{w(x_j)+w(x_i)} \cdot x_j \text{ for all } i \in N$$

(with the convention $\frac{0}{0}=1$ in the equation above)

Proof.
Because the distribution of each variable Z_i is non atomic, and these variables are independent, the probability of a tie $Z_i = Z_j$ is null. Hence, with the notations above, the probability that ordering σ is selected is that of the event $\{Z_1 \leq Z_2 \leq \cdot\cdot \leq Z_n\}$, where Z_i is Weibull with parameter w_i. This is precisely (4), as follows from a simple computation that we reproduce for the case $n=3$:

$$\int_0^1 w_3 \cdot z^{w_3-1}\{\int_0^z w_2 \cdot t^{w_2-1}\{\int_0^t w_1 \cdot s^{w_1-1} ds\}dt\}dz =$$

$$\int_0^1 w_3 \cdot z^{w_3-1}\{\int_0^z w_2 \cdot t^{w_1+w_2-1} dt\}dz = \frac{w_2}{w_1+w_2}\int_0^1 w_3 \cdot z^{w_1+w_2+w_3} dz$$

There is an alternative, perhaps more intuitive definition of quasi-proportional rules, for those problems (N,x) where the numbers $w(x_i)$ are all rationals, hence $w(x_i)=\frac{b_i}{d}$ for some integers b_i, d. For each agent i put b_i balls of colour i in an urn and empty the urn by successive draws with uniform probability and without replacement; schedule the jobs in the order in which each colour vanishes in the urn (when the last ball of this colour is drawn). This generates precisely the random ordering in Lemma 2[5].

The Uniform rule is quasi-proportional with $w(a) = 1$ for all $a > 0$. The Shortest (resp. Longest) Job First rule is the limit of the quasi-proportional rules $w(a) = a^\alpha$, when α goes to $+\infty$ (resp. to $-\infty$). The quasi-proportional rule to which our main results (Section 7) are devoted is the *Proportional* rule, for which $w(a) = a$. Another remarkable rule (see Lemma 7 in Section 6.2) is the *Quadratic* rule for which $w(a) = a^2$.

Finally we note an important property of quasi-proportional rules revealed by equation (4): the probability p_σ is computed recursively by finding first the probability $\frac{w_n}{\sum_1^n w_k}$ that agent $\sigma^{-1}(n)$ be ranked last in problem (N, x), multiplying this number by the probability that agent $\sigma^{-1}(n-1)$ be ranked last in the reduced problem $(N \diagdown \{\sigma^{-1}(n)\}, x_{[N \diagdown \sigma^{-1}(n)]})$, and so on. In other words the scheduling rule is entirely determined once we know the probability distribution of the job served last at every problem.

Definition 6

The scheduling rule ρ is *recursive* if there exists for all N a function π from $\mathbb{R}^N_{++}$ into $\Delta(N)$, such that for all N, x, and $\sigma \in \Phi(N)$ with $\sigma^{-1}(n) = i$, we have

$$p_\sigma(N, x) = \pi_i(N, x) \cdot p_{\sigma[N \diagdown i]}(N \diagdown i, x_{[N \diagdown i]})$$

We can extend the definition of $\pi(N, \cdot)$ to $\mathbb{R}^N_+ \diagdown \{0\}$ by setting $\pi_i(N, x) = 0$ whenever $x_i = 0$. Then the equation above implies efficiency.

The serial rule is an example of a parametric rule that is *not* recursive. Take the problem $N = \{1, 2, 3\}, x = (1, 2, 3)$ and check:

$$prob\{Z_1 < Z_2 < Z_3\} = \frac{19}{36};\ prob\{\max\{Z_1, Z_2\} < Z_3\} = \frac{11}{18};\ prob\{Z_1 < Z_2\} = \frac{3}{4}$$

Our next result, of which the proof is relegated to the Appendix, explains the central role of quasi-proportional rules in our model. Recall that a rule ρ (resp. a method μ) is *anonymous* if the mapping $x \to p(N, x)$ (resp. $x \to y(N, x)$) is symmetric in all variables on $\mathbb{R}^N_{++}$ (resp. $\mathbb{R}^N_+$).

Proposition 1

A scheduling rule is quasi-proportional if and only if it is separable, recursive, anonymous and meets the following ***Positivity*** *property*

[5] I am grateful to R.J. Aumann for pointing out this interpretation.

$$prob\{\sigma(j) < \sigma(i)|x\} > 0 \text{ for all } N, x, x \gg 0, \text{ and all } i, j.$$

Note that Positivity is needed in the above statement: Shortest Jobs First and Longest Jobs First meet all three other properties. It is not hard to describe the more complicated set of separable and recursive scheduling rules, by mimicking the proof techniques of Theorem 2 in Moulin and Stong [2002]. We omit the details for brevity.

Remark 1

An dual definition of Recursivity works by drawing the agent scheduled *first* in the given problem, then repeating in the reduced problem and so on. For instance if we schedule i first with probability $\frac{1}{w(x_i)}$, the resulting rule is separable and the probability that job j precedes job i is $\frac{w(x_i)}{w(x_j)+w(x_i)}$, just like with the rule of Definition 5; these two rules are different with three or more agents, but they implement the same method. However this dual definition of Recursivity does not yield interesting Split-proof rules. For instance the dual Proportional rule (scheduling i first with probability $\frac{1}{x_i}$ is not Split-proof.

5 Merging and Splittinig

5.1 Merge-proofness and demand monotonicity

We start with a 5 agents example. In the problem (N, x), where $N = \{1,2,3,4,5\}$, asssume that the coalition $\{1,2,3\}$ *merges* its jobs. This means that they report as a single agent of the coalition, say 1^*, a job of length $x^*_{1^*} = x_1 + x_2 + x_3$, and choose freely a scheduling order, possibly a randomized one, of their "true" jobs in the time interval chosen by the server to process the merged job. Write $y^*_{1^*}$ for the expected wait of the merged job of size $x^*_{1^*}$, so that $y^*_{1^*} - x^*_{1^*}$ is the expected length of time during which job $x^*_{1^*}$ will wait for the jobs of agents 4 and 5. The latter expected wait is borne by all three agents upon merging. The other part of their true wait comes from the processing order of jobs $1, 2$, and 3: Lemma 1 describes the range $F(\{1,2,3\},(x_1,x_2,x_3))$ of such profiles of expected waits. Summing up we see that the profile $\widetilde{y} = (\widetilde{y}_1, \widetilde{y}_2, \widetilde{y}_3)$ of expected waits is feasible for $\{1,2,3\}$ after merging if and only if $\widetilde{y} \in (y^*_{1^*} - x^*_{1^*}) \cdot (1,1,1) + F(\{1,2,3\},(x_1,x_2,x_3))$. If y is the intial profile of waits for our three agents, merging is profitable if and only if they can find $\widetilde{y}$ such that $\widetilde{y} \le y$ with at least one strict inequality.

Observe that the description of successful merges only uses the scheduling *method*, not the actual scheduling *rule* from which it is derived. This simplifies greatly the analysis of the Merge-proofness property, of which the definition is a straightforward generalization of the above example. We introduce some notations first. Given a scheduling problem (N, x), a proper subset (coalition) T of N, and an agent $i^* \in T$, the merger of T into i^* creates a new problem (N^*, x^*), where $N^* = (N \diagdown T) \cup \{i^*\}$, $x^*_{i^*} = \sum_T x_i$, $x^*_j = x_j$ for $j \in N \diagdown T$. We write e for the vector in $\mathbb{R}^N$ with all coordinates equal to 1. Finally, as in Lemma 1, the set $F(T, x_{[T]})$ consists of the profiles of expected waits feasible for coalition T when it is scheduled before $N \diagdown T$.

Definition 7

Given a scheduling method μ, a problem $(N, x), x \geq 0$, and T, i^* as above, we write $y = \mu(N, x)$ for the expected waits before merging and $y^* = \mu(N^*, x^*)$ for the (reported) waits after merging. We say that μ is vulnerable to merging by T, i^* at problem (N, x) if there exists a vector $\widetilde{y}_{[T]} \in \mathbb{R}^T$ such that

$$\widetilde{y}_{[T]} \in (y^*_{i^*} - \sum_T x_i) \cdot e_{[T]} + F(T, x_{[T]}); \, \widetilde{y}_{[T]} \leq y_{[T]} \text{ and } \widetilde{y}_{[T]} \neq y_{[T]} \qquad (5)$$

We say that the method μ is *merge-proof* if it is not vulnerable to merging at any problem by any coalition. We say that the scheduling rule ρ is merge-proof if the associated method is.

A null agent, $x_i = 0$, does not wait in the original problem, but may wait after merging. On the onther hand, adding a null agent to a merging coalition never helps those agents. Threfore in checking merge-proofness of a method, it is enough to check property (5) for coalitions T such that $x_{[T]} \gg 0$.

Our contention that Merge-proofness is "easy" to achieve relies on the mild responsiveness property stating that an agent's share of the delay externality be non decreasing in the size of his own job.

Definition 8

The scheduling method μ is *demand monotonic* if for all N, and $i \in N$, $y_i(N, x) - x_i$ is non decreasing in x_i. The scheduling rule ρ is demand monotonic if the associated method is.

The dual interpretation of this property, both as a normative and an incentive-compatibility statement, is standard (see e.g., Moulin and Sprumont [2003]): increasing the size of my job augments the delay externality, and it is only fair that my share of this externality should not decrease; on the other hand if the property fails, I would benefit by an artificial increase of the size of my job, provided I am able to enjoy the service as soon as my real job is completed.

Lemma 3

i) The θ-separable scheduling method (Definition 3) is demand monotonic if and only if for all i, j, all $b > 0$, $\theta^{i,j}(a, b)$ is non decreasing in a;
ii) The parametric rule $\{F_a, a > 0\}$ is demand monotonic if $a \to Z_a$ is stochastic-dominance monotonic , i.e., $\{a \leq b\} \Rightarrow \{F_b(z) \leq F_a(z)$ for all $z\}$.
iii) The $w-$quasi-proportional rule is demand monotonic if and only if w is non decreasing.

Proof of statements i, iii.

If $\theta^{i,j}(a, b)$ is non decreasing in a, equation (2) implies Demand Monotonicity at once. Conversely, if the $\theta-$separable method is demand monotonic, $y_i(\{i, j\}, x) - x_i = \theta^{i,j}(x_i, x_j) \cdot x_j$ is non decreasing in x_i, completing the proof of statement i. Statement iii follows at once.

The proof of statement ii is in the Appendix. Here is a demand monotonic parametric rule for which $a \to Z_a$ is not stochastic-dominance monotonic. For $a < 1$, let $Z_a = 2$ with probability one, and for $a \geq 1$, let $Z_a = 1$ or 3 with respective probabilities $\frac{1}{4}$ and $\frac{3}{4}$.

Shortest Jobs First, the Serial rule (Example 1), the Uniform[6], Proportional, and Quadratic rules are all demand monotonic. Longest Jobs First is emphatically *not* demand monotonic.

Proposition 2

An anonymous and demand monotonic separable scheduling method is merge-proof. For instance a demand monotonic parametric rule is merge-proof.

Proof

We show first a preliminary result.The scheduling method μ is merge-proof if for every $(N, x), x \geq 0, T, i^*$ as in Definition 7, with $y = \mu(N, x)$ and $y^* = \mu(N^*, x^*)$, we have

$$\sum_T x_i \cdot y_i + \sum_{T(2)} x_i \cdot x_j \leq (\sum_T x_i) \cdot y^*_{i^*} \tag{6}$$

Recall that $T(2)$ is the set of non ordered pairs in T. Consider a vector $\widetilde{y}_{[T]}$ in the set defined in property (5). Using the notation $\sum_T x_i = x_T$, we compute from Lemma 1

[6]Note that for the Uniform rule, $y_i - x_i$ is actually independent of x_i. It is easy to see that the Uniform *method* is characterized by this property and Anonymity.

$$\sum_T x_i \cdot \widetilde{y}_i = x_T \cdot (y^*_{i^*} - x_T) + v(T,x) = x_T \cdot y^*_{i^*} - \sum_{T(2)} x_i \cdot x_j$$

Thus (6) is equivalent to $\sum_T x_i \cdot y_i \leq \sum_T x_i \cdot \widetilde{y}_i$, and $\{\widetilde{y}_{[T]} \leq y_{[T]}$ and $\widetilde{y}_{[T]} \neq y_{[T]}\}$ is impossible.

We pick now an anonymous $\theta-$separable method and prove (6) for any N, x, T, i^*. We develop (6) using (2). The term $\sum_T x_i \cdot y_i$ in the LHS is computed as

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

Thus the LHS of (6) is $x_T^2 + \sum_{i \in T, j \in N \setminus T} \theta(x_i, x_j) \cdot x_i \cdot x_j$. The RHS is $x_T \cdot [x_T + \sum_{j \in N \setminus T} \theta(x_T, x_j) \cdot x_j]$, therefore (6) amounts to

$$\sum_{j \in N \setminus T} x_j \cdot [\sum_{i \in T} \theta(x_i, x_j) \cdot x_i] \leq \sum_{j \in N \setminus T} x_j \cdot [\theta(x_T, x_j) \cdot x_T].$$

This holds if $a \to \theta(a,b) \cdot a$ is superadditive for $b > 0$ (the inequality above always holds if $x_j = 0$). As θ is non negative, the latter is true if $\theta(a,b)$ is non decreasing in a (because θ is non negative). The proof of Proposition 2 is complete.

A plausible statement is that an anonymous separable method is merge-proof *if and only if* it is demand monotonic. This is not true, however. We saw in the above proof that the anonymous $\theta-$separable method is merge-proof *if* $a \to \theta(a,b) \cdot a$ is superadditive. Hence the $w-$quasi-proportional rule is merge-proof if $a \to a \cdot w(a)$ is superadditive[7]. We can choose such a function w that fails to be non decreasing, and the corresponding quasi-proportional rule will not be demand monotonic (Lemma 3).

We observe finally that the Uniform method meets a stronger property than merge-proofness, namely it is *Merge-invariant.* This means that when a coalition merges, it can achieve after merging precisely the same profile of expected wait as before merging, and no better. Formally Merge-invariance requires that for all N, x, T, i^* as in the premises of Definition 7,

[7]Indeed this implies for all $a, a' : w(a + a') \geq \frac{a}{a+a'} w(a) + \frac{a'}{a+a'} w(a')$; applying the concave and increasing function $f(z) = \frac{z}{z+w(b)}$ to both sides of this inequality gives the superadditivity of $a \cdot \theta(a,b) = \frac{a \cdot w(a)}{w(a)+w(b)}$ as claimed.

$$y_{[T]} \in (y^*_{i^*} - \sum_T x_i) \cdot e_{[T]} + F(\widetilde{T}, x_{[T]}). \tag{7}$$

To check this for the uniform method, set $z_{[T]} = y_{[T]} - (y^*_{i^*} - \sum_T x_i) \cdot e_{[T]}$, and compute $z_i = (x_i + \frac{1}{2}x_{N\diagdown i}) - \frac{1}{2}x_{N\diagdown T} = x_i + \frac{1}{2}x_{T\diagdown i}$ for all $i \in T$. By Lemma 1, $z_{[T]} \in F(T, x_{[T]})$ as desired. In fact we can say more.

Proposition 3

The Uniform method is the only merge-invariant scheduling method.

Note that this characterization holds within the full space of scheduling methods, separable or not. In view of the dreadful normative features of the Uniform rule discussed in Section 6, we interpret Proposition 3 as a negative result: merge-invariance is not compatible with even a modicum of responsiveness and limited liability.

Proof

Given $N, S, S \subseteq N$, the set of real valued functions $x \to f(x)$ defined on $\mathbb{R}^N_+$, that depend only upon $x_{[N\diagdown S]}$ and $x_S = \sum_S x_i$ is denoted $\Lambda(S)$. Fix a merge-invariant method μ, and N, x, T, i^* as in the premises of Definition 7. Property (7) implies

$$\begin{aligned} \sum_T x_i \cdot y_i &= x_T \cdot (y^*_{i^*} - x_T) + v(T, x) \\ &\Rightarrow \sum_T x_i \cdot y_i + \sum_{T(2)} x_i \cdot y_i = x_T \cdot y^*_{i^*} \in \Lambda(T) \end{aligned}$$

Setting $z_i = y_i - \frac{1}{2}x_i$ this gives $\sum_T x_i \cdot z_i \in \Lambda(T)$. On the other hand $y \in F(N, x)$ implies $\sum_N x_i \cdot y_i = \sum_T x_i \cdot (z_i + \frac{1}{2}x_i) = v(N, x)$, from which we derive $\sum_N x_i \cdot z_i = \frac{1}{2}x_N^2 \in \Lambda(N)$. As the choice of T wass arbitrary, we conclude $\sum_T x_i \cdot z_i \in \Lambda(T) \cap \Lambda(N\diagdown T)$ for all T, and in particular $x_i \cdot z_i = \varphi_i(x_i, x_N)$ for some function φ_i. Now from the equality $\sum_i \varphi_i(x_i, x_N) = \frac{1}{2}x_N^2$, the inequality $\varphi_i \geq 0$, and $\varphi_i(0, x_N) = 0$, we get $z_i = \frac{1}{2}x_N$ by a straightforward application of the Cauchy equation (Aczel [1966]).

5.2 Split-proofness

We start by a three agents example. In the problem (N, x) where $N = \{1, 2, 3\}$, assume that agent 3 reports as two agents $3, 4$ and *splits* her (true) job of size x_3 into two jobs of sizes x_{*3}, x_{*4}, with $x_{*3} + x_{*4} = x_3$. After splitting, agent 3 must wait until both "sub-jobs" are completed hence the delay she experiences from jobs 1 and 2 is evaluated by computing the probabilities that these jobs are scheduled before at least one of the sub-jobs. Setting $x_* = (x_1, x_2,\ x_{*3},\ x_{*4})$ and using the notations of Definition 1, we see that the split benefits agent 3 iff

$$\sum_{i=1,2} prob\{\sigma(i) < \max(\sigma(3), \sigma(4)) | x_*\} \cdot x_i < \sum_{i=1,2} prob\{\sigma(i) < \sigma(3) | x\} \cdot x_i$$

The following notations are used in the general definition of Split-proofness. Given a problem $(N, x), x \gg 0$, a coalition T such that $T \cap N = \varnothing$, and an agent $i_* \in N$, the splitting of i_*into $i_* \cup T$ creates a new problem $(N_*, x_*), x_* \geq 0$, where $N_* = N \cup T$, $x_{i_*} = \sum_{i_* \cup T} (x_*)_i$, $(x_*)_j = x_j$ for $j \in N \diagdown i^*$. Note that the initial job is positive, $x_{i_*} > 0$, because a null job has no incentive to deviate; on the other hand we allow some coordinate of x_* to be null, because by splitting x_{i_*} into x^* such that $(x_*)_{i_*} = 0, (x_*)_j = x_{i_*}$, agent i_* effectively assumes a new identity as agent j and such maneuvers are both realistic and important in our model (see the proof of Lemma 4 below).

Definition 9

Fix a scheduling rule ρ, a problem $(N, x), x \gg 0$, and T, i_* as above. We say that ρ is *split-proof* at (N, x) with respect to T, i_* if

$$y_{i_*}(N, x) \leq x_{i_*} + \sum_{j \in N \diagdown i_*} prob\{\sigma(j) < \max_{i_* \cup T} \sigma(i) | x_*\} \cdot x_j \tag{8}$$

We say that ρ is split-proof if it is split-proof for all (N, x) and T, i_*.

The right-hand side of the above inequality is the (true) expected wait of agent i_* after the split. Note that the definition of Split-proofness applies to a scheduling rule: knowledge of the method associated with a given rule is not enough to decide whether or not the rule is split-proof.

For a separable scheduling rule (Definition 2), Split-proofness takes a much simpler form. Recall our notation $\theta^{i,j}(x_i, x_j)$ for the probability that agent j's job of size x_j precedes agent i's job of size x_i.

Lemma 4

The separable scheduling rule ρ is split-proof if and only if the corresponding method is anonymous ($\theta^{i,j} = \theta$ is independent of i, j), and moreover for all $S, i, i \notin S$, and all positive numbers $b, a_j, j \in S$,

$$\theta(b, \sum_S a_j) \geq prob\{\max_S \sigma(j) < \sigma(i)|(b, a_j)\} \quad (9)$$

Proof

Statement "if". Fix (N, x) and T, i_* as in the premises of Definition 9. By assumption $y_{i_*}(N, x) - x_{i_*} = \sum_{N \setminus i_*} \theta(x_{i_*}, x_j) \cdot x_j$. Next the term $prob\{\sigma(j) < \max_{i_* \cup T} \sigma(i)|x_*\}$ depends only upon $(x_*)_{[j \cup i_* \cup T]}$ by Separability of ρ. Thus inequality (8) follows from $\theta(x_{i_*}, x_j) \leq prob\{\sigma(j) < \max_{i_* \cup T} \sigma(i)|x_*\}$, which is precisely (9).

Statement "only if". We show first that $\theta^{i,j}$ is independent of i, j. Fix $a, b > 0$ and consider in problem $N = \{1, 2\}, x = (a, b)$, the split of agent 1 into agents $1, 3$ with $(x_*)_1 = 0, (x_*)_3 = a$. In the split problem, agent 1 is scheduled first by efficiency, so the probability that agent 2 is last is just $\theta^{2,3}(b, a)$. Thus split-proofness implies $\theta^{2,1}(b, a) \geq \theta^{2,3}(b, a)$. Exchanging the roles of $1, 3$ gives $\theta^{2,1} = \theta^{2,3}$. By $\theta^{i,j}(b, a) + \theta^{j,i}(a, b) = 1$ we have now $\theta^{1,2} = \theta^{3,2}$; Anonymity follows.

Next fix $S, i, i \notin S$,and b, a_j as in the premises of (9) and choose an agent $1 \in S$. In the problem $N = \{1, i\}, x = (a_S, b)$ a split by 1 to S with $(x_*)_j = a_j$ for all $j \in S$ cannot benefit agent 1: this gives inequality (9), and completes the proof of Lemma 4.

Examples of split-proof separable rules include the Uniform and Proportional rules (by Proposition 4), and Longest Job First. On the other hand neither the serial rule (Example 1) nor the Quadratic rule is split-proof. The latter claim follows from the Corollary to proposition 4 below. To check the former one, recall that Z_a is uniform on $[0, a]$, and consider $S = \{1, 2\}, i = 3$ and $b = 3, a_1 = a_2 = 1$. Then compute $\theta(3, 2) = \frac{2}{3} < prob\{\max(\sigma(1), \sigma(2)) < \sigma(3)|(3, 1, 1)\} = \frac{7}{9}$.

Applying Lemma 4 to parametric rules, we find a very convenient sufficient condition for Split-proofness in that set of rules. Fix a parametric rule $\{F_a, a > 0\}$ and assume for a moment that all distributions F_a are atomless (i.e., F_a is continuous). As ties occur with probability zero, the inequality (9) amounts to

$$\begin{aligned} prob\{Z_{a_S} &\leq Z_b\} \geq prob\{\max_S Z_{a_j} \leq Z_b\} \Longleftrightarrow \\ \int_0^\infty F_{a_S}(z) \cdot dF_b(z) &\geq \int_0^\infty \Pi_{j \in S} F_{a_j}(z) \cdot dF_b(z) \end{aligned}$$

which holds true if $F_{a_S}(z) \geq \Pi_{j \in S} F_{a_j}(z)$. The latter inequality is in fact sufficient for Split-proofness, whether or not the corresponding distributions have atoms.

Proposition 4

If the cdfs $\{F_a, a > 0\}$ satisfy $F_a \cdot F_b \leq F_{a+b}$ for all $a, b > 0$, the corresponding parametric rule is split-proof. If they satisfy $F_a \cdot F_b \leq F_{a+b} \leq F_a$ for all $a, b > 0$, the rule is both merge-proof and split-proof.

Corollary

The $w-$quasi-proportional rule is split-proof if and only if w is subadditive. It is merge-proof and split-proof if w is non decreasing and subadditive.

The proof of proposition 4 in the case of distributions with atoms, is relegated to the Appendix.

Proof of Corollary. For the $w-$quasi-proportional rule, the cdfs F_a are atomless. Inequality $F_a(z) \cdot F_b(z) \leq F_{a+b}(z)$ is always true if $z \geq 1$, and amounts to $w(a) + w(b) \geq w(a+b)$ if $z < 1$. This proves the "if" statement. Conversely, suppose the $w-$rule is split-proof. By Separability, inequality (9) must be true. As $\Pi_{j \in S} F_{a_j}(z) = z^{\sum_S w(a_j)}$, the right-hand side in (9) equals $\frac{w(b)}{w(b)+\sum_S w(a_j)}$, therefore this inequality reads

$$\theta(b, \sum_S a_j) = \frac{w(b)}{w(b) + w(\sum_S a_j)} \geq \frac{w(b)}{w(b) + \sum_S w(a_j)}$$

implying the subadditivity of w.

The Corollary identifies a rich class of merge-proof and split-proof quasi-proportional rules. Suppose we restrict attention to *scale invariant* quasi-proportional rules, namely such that the scheduling order follows the same distribution in problem (N, x) and in problem $(N, \lambda \cdot x)$, for any positive scaling factor λ. It is easy to check with the help of lemma 2[8] that this forces to choose a weight function of the form $w(a) = a^\alpha$, for some real number α. Among these rules, Merge-proofness imposes $\alpha \geq 0$, and Split-proofness imposes $\alpha \leq 1$. The Uniform rule, $\alpha = 0$, and the Proportional rule, $\alpha = 1$, are the two extreme points of this interval of rules.

But we stress that there are many other merge-proof and split-proof parametric scheduling rules. Two examples are the cdfs $F_a(z) = \min\{\frac{(1+a)z}{a+z}, 1\}$, and $F_a(z) = e^{-\frac{a^2}{z(a+z)}}$. In both cases one checks that $F_a(z)$ is non increasing in a and $F_a(z) \cdot F_b(z) \leq F_{a+b}(z)$. Moreover, Merge-proofness and Split-

[8] And a standard application of the Cauchy equation.

proofness are stable by convex combinations of rules, i.e., if ρ and ρ' are two merge-proof (rersp. split-proof) scheduling rules , for any $\lambda \in [0,1]$ the rule $(N,x) \to \lambda \cdot p(N,x) + (1-\lambda) \cdot p'(N,x)$ is merge-proof (resp.split-proof) too. Notice that parametric rules are *not* stable by convex combinations, thus we generate many more merge-proof and split-proof rules in this fashion[9].

The Uniform and Proportional rules are the two extreme points of the one-dimensional interval of scale invariant quasi-proportional rules that are merge-proof and split-proof. They also stand out for two other, more fundamental reasons: the Uniform rule is merge-invariant, whereas the Proportional rule is split-invariant. To check the latter claim, observe that the cdfs $F_a(z) = z^a$ satisfy $F_a \cdot F_b = F_{a+b}$ for all $a, b > 0$, therefore inequality (9) is actually an equality for this rule, and the same holds true for inequality (8), expressing that the expected wait of agent i_* is the same before and after the split.

While Merge- invariance single-handedly characterizes the Uniform rule (Proposition 3), the Proportional rule is not the only split-invariant rule. However we show in section 7 that the Split-invariance property characterizes the Proportional rule both within the class of separable rules (Theorem 1) and within that of recursive rules (Theorem 2).

We conclude this Section with another sufficient condition for Split-proofness, generalizing the subadditivity of w to arbitrary recursive rules (Definition 6).

Lemma 5
The recursive scheduling rule ρ is split-proof if for all $N, i \in N, j \notin N$, and all $x \in \mathbb{R}_+^N, \widetilde{x} \in \mathbb{R}_+^{N \cup j}$

$$\{x_{[N \diagdown i]} = \widetilde{x}_{[N \diagdown i]},\, x_i = \widetilde{x}_i + \widetilde{x}_j\} \Longrightarrow \{\pi_k(N,x) \geq \pi_k(N \cup j, \widetilde{x}) \text{ for all } k \in N \diagdown i\}.$$

This says that when agent i splits his job in two (or more) pieces, the probability that another job is scheduled last does not become larger. Loosely speaking, each function $\pi_k(N,x)$ is subadditive upon splitting. The proof is in the Appendix.

[9]Separability is stable by convex combinations, but Recursivity is not.

6 Responsiveness and Liability

6.1 The Ranking axiom

We discuss now two normative properties of scheduling methods, Ranking and Limited Liability. We show in the next section that within split-proof rules, these properties clearly point toward the Proportional rule.

Definition 9

The scheduling method μ meets *Ranking* if for all $N, x, x \geq 0$, and $i, j \in N$

$$x_i \leq x_j \Rightarrow y_i(N,x) - x_i \leq y_j(N,x) - x_j$$

As discussed in section 1, the interpretation of Ranking is similar to that of Demand Monotonicity, with the difference that it makes interpersonal comparison of delay shares. Another similarity is that both axioms are esay to express for anonymous separable methods, hence for all parametric rules.

Lemma 6

The anonymous $\theta-$separable method ($\theta^{i,j} = \theta$ does not depend on i, j) meets Ranking if and only if it is demand monotonic ($\theta(a,b)$ is non decreasing in a for all $b > 0$) and moreover

$$\text{for all } i, j, \text{ all } a, b > 0 : \theta(a,b) \leq \frac{a}{a+b}.$$

Proof

The "if" statement follows by developing the inequality $y_i - x_i \leq y_j - x_j$ with the help of (2). To prove "only if", apply first Ranking in the two person problem $N = \{1,2\}, x = (a,b)$, to get the upper bound on $\theta(a,b)$. Next we show by contradiction that $\theta(a,b)$ is non decreasing in a. Suppose for some $a, a', b > 0$, we have $a < a'$ yet $\theta(a,b) > \theta(a',b)$. Consider the problem $N = \{1,2,..,n\}, x = (a,a',b,..,b)$: for n large enough we get $y_1 - x_1 > y_2 - x_2$, in contradiction of Ranking.

For instance the Serial method (Example 1) meets Ranking.

Comparing statement $i)$ in Lemma 3 and Lemma 6, we see that Ranking is strictly more demanding than Demand Monotonicity for split-proof separable methods.

Another consequence of Lemma 6 is that the $w-$quasi-proportional method meets Ranking *if and only if* $\frac{w(a)}{a}$ is non decreasing. E.g., for $w(a) = a^{\alpha}$, Ranking holds if and only if $\alpha \geq 1$. Hence a tension between Ranking and

Split-proofness: the $w-$quasi-proportional rule is split-proof iff w is subadditive (Corollary to Proposition 4); it meets Ranking iff $\frac{w(a)}{a}$ is non decreasing, which implies that w is superadditive. Both properties are true iff w is linear, namely when the rule is the Proportional one. As explained in the next Section, this tension occurs in the much larger classes of separable or recursive scheduling rules.

6.2 Limited liability

Definition 10
The *liability* of a scheduling method μ is the function $\lambda(n,a) = \sup_{i,x_{-i}} y_i(a, x_{-i})$, where n is an integer no less than 2 and $a > 0$, and where the supremum is taken over all subsets N of $\mathcal{N}$ with n agents, all choices of $i \in N$, and all job profiles $x_{-i} \in \mathbb{R}_+^{N \diagdown i}$ for the other agents.

The liability measures the worst conceivable expected wait for an agent who only knows the number of other users, and the size of her own job. To limit the liability is of paramount importance when non participation is an option, for instance if another server, slower but with a low liability, is available.

For the Uniform rule, $y_i(a, x_{-i}) = a + \frac{1}{2} x_{N \diagdown i}$ hence $\lambda(n,a) = +\infty$. Together with the violation of Ranking, this is a damning critique of this rule. By contrast Shortest Jobs First has $\lambda(n,a) = n \cdot a$, and so does the Proportional rule. The former claim is clear, and the latter results from the following computation

$$\lambda(n,a) = \sup_{x_{-i} \geq 0} \{a + \sum_{N \diagdown i} \frac{a}{a + x_j} \cdot x_j\} = a + \sum_{N \diagdown i} \sup_{x_j} \frac{a}{a + x_j} \cdot x_j = n \cdot a$$

This level of liability is about twice larger than the minimal feasible liability:

Lemma 7
i)For any scheduling method μ, we have $\lambda(n,a) \geq \frac{n+1}{2} \cdot a = \lambda^*(n,a)$.
ii) The $\theta-$separable method has the minimal liability λ^* if and only if $\theta^{i,j}(a,b) \leq \frac{a}{2b}$ for all i, j, and all $a, b > 0$. Examples include the Serial method (Example 1) and the Quadratic method ($w(a) = a^2$).

Proof
For an arbitrary method μ, consider a problem (N, x) where $x_i = a$ for all $i \in N$. Feasibility of $y = \mu(N, x)$ (Lemma 1) implies $a \cdot (\sum_N y_i) = v(N, x) =$

$\frac{n(n+1)}{2} \cdot a^2$. Therefore $\max_N y_i \geq \frac{n(n+1)}{2} \cdot a$, establishing statement i. Next assume that μ is separable. For any N, i, and $a > 0$, compute the worst expected wait from equation (2):

$$\sup_{x_{-i}} y_i(a, x_{-i}) = a + \sum_{j \in N \diagdown i} \sup_{x_j} \theta^{i,j}(a, x_j) \cdot x_j \tag{10}$$

from which statement ii is straightforward.

The goal of minimizing the liability of a rule conflicts with that of ensuring it is split-proof; the simplest way to demonstrate this important trade-off is to look at the line of scale invariant quasi-proportional rules, namely $w(a) = a^\alpha$. Using equation (10) it is easy to compute

- if $\alpha < 1$, then $\lambda(n, a) = +\infty$ for all n, a;
- if $1 \leq \alpha < +\infty$, then $\lambda(n, a) = (1 + (n-1)\frac{(\alpha-1)^{1-\frac{1}{\alpha}}}{\alpha}) \cdot a$; thus for $1 \leq \alpha \leq 2$, $\lambda(n, a)$ decreases from $n \cdot a$ to $\frac{n+1}{2} \cdot a$, and it increases from $\frac{n+1}{2} \cdot a$ to $n \cdot a$ for $2 \leq \alpha < +\infty$.

Recall that a rule in this family is splitproof iff $\alpha \leq 1$: therefore the Proportional rule is the only split-proof rule with finite liability. This observation generalizes to quasi-proportional rules in the following way:

- if $\frac{w(a)}{a}$ is non increasing, and $\lim_{+\infty} \frac{w(a)}{a} = 0$, and/or $\lim_0 \frac{w(a)}{a} = +\infty$, then $\lambda(n, a) = +\infty$ for all n, a;
- if $\frac{w(a)}{a}$ is non decreasing, then $\lambda(n, a) \leq n \cdot a$ for all n, a.

On the other hand the $w-$rule is split-proof when $\frac{w(a)}{a}$ is non increasing (by the Corollary to Proposition 4). The proof of the two claims above is again a straightforward consequence of (10). We omit the details.

7 Characterizations of the proportional rule and method

We offer first a characterization of the proportional *method* under Separability, then we characterize the proportional *rule* in the class of recursive rules.

Our first result explains the general trade-off between Ranking and Split-proofness, its relation to the liability measure, and the critical role of the Proportional rule: under separability, the liability of a split-proof rule is not smaller – and that of a rule meeting Ranking is not larger – than that of the Proportional rule.

Lemma 8

i)If the separable scheduling rule ρ is split-proof, then $\lambda(n,a) \geq n \cdot a$ for all n, a;

ii)If the separable scheduling method μ meets Ranking, then $\lambda(n,a) \leq n \cdot a$ for all n, a.

Proof

Statement i) is a consequence of Lemma 9 in the Appendix, as explained there. For statement ii), consider the $\theta-$separable method; Ranking implies for all i, j and all $a, b > 0$ such that $a \leq b$:

$$\begin{aligned} y_i(\{i,j\},(a,b)) - a &\leq y_j(\{i,j\},(a,b)) - b \Longleftrightarrow \theta^{i,j}(a,b) \cdot b \leq \theta^{j,i}(b,a) \cdot a \\ &\Longleftrightarrow \theta^{i,j}(a,b) \leq \frac{a}{a+b} \leq \frac{a}{b} \end{aligned}$$

and $\lambda(n,a) \leq n \cdot a$ then follows from (10).

Theorem 1

i)If the separable scheduling rule ρ is split-invariant, then its method is Proportional: $y_i(N,x) = x_i + \sum_{N \diagdown i} \frac{x_i \cdot x_j}{x_i + x_j}$ for all $N, x, x \geq 0$, and i.

ii)If the separable rule ρ is split-proof and meets one of Ranking or $\{$for some n, $\lambda(n,a) \leq n \cdot a$ for all $a\}$, then its method is Proportional.

Theorem 2

i) The Proportional rule is the only recursive and split-invariant scheduling rule.

ii) The Proportional rule is the only recursive and split-proof rule that also meets one of Ranking or $\{\lambda(2,a) \leq 2a$ for all $a\}$.

Both Theorems are proven in the Appendix by essentially the same argument. Yet in the class of separable rules, we only characterize the Proportional method, whereas in the recursive class we capture the rule itself. For $|\mathcal{N}| = 4$ it is easy to construct a separable and split-invariant rule that is not the Proportional one. I conjecture that such a construction is possible for $|\mathcal{N}| = \infty$ as well. This suggests that Theorem 1 cannot be improved into a characterization of the Proportional rule.

Remark 2
In the class of parametric rules (Definition 4), a variant of Theorem 2 holds true. Recall that the parametric class neither contains nor is contained in the recursive class. If a rule is split-invariant, or if it is split-proof and meets either Ranking or $\{\lambda(2,a) \leq 2a$ for all $a\}$, the argument in Sections 9.5,9.6 show that for all N, x, and all i, the probability that i is scheduled last is proportional to x_i. I have shown that this property characterizes the Proportional rule among *continuous* (atomless) parametric rules[10], and conjecture that the same holds among all parametric rules.

8 Open problems and future research

1. Parametric rules have a natural extension to queuing problems, where jobs are born at arbitrary dates. Each time a new job i of size x_i is born, we draw the variable Z_i with cdf F_{x_i} independently of the draws of all Z_j corresponding to jobs born earlier, and we process jobs in the preemptive priority defined by these realizations. This definition preserves the property of merge-proofness provided the reported merged job is born not earlier than the youngest of the component jobs; it preserves split-proofness as well, if the split jobs are born not earlier than the true job. Thus the extended Proportional rule remains equally appealing because of these two properties, plus Ranking and the same liability cap $n \cdot a$, where n is the number of jobs not older than one's own. Obviously the definition of the liability should be modified if the arrival of new jobs is Poisson with a known arrival rate. Whether or not the characterizations of the Proportional extend to the queuing context is left for future research.

2. Minimizing total waiting time is not compatible with Split-proofness, so a natural question is to understand the trade-off between these two requirements. The minimal waiting time in problem (N, x) is $w(N, x) = \sum_N x_i + \sum_{N(2)} \min\{x_i, x_j\}$. Define the index of "excess wait" $ew(n)$ for a method μ as the largest ratio $\frac{\sum_N y_i}{w(N,x)}$ over all problems. An interesting question is to evaluate the minimal index of excess wait among split-proof (and merge-proof) scheduling rules. A plausible conjecture is that the Proportional rule achieves this minimum. Yet it is hard to compute the index of excess wait of a specific rule such as the Proportional one, therefore the

[10] The proof is available upon request.

conjecture is not an easy one to crack.

Another interesting trade-off is between the utilitarian goal to minimize the index of excess wait, and that of guaranteeing to each participant the lowest feasible liability. We saw in Section 6 that Shortest Jobs First offers a liability about twice as large as the lowest feasible one. Conversely, the Quadratic rule guarantees the lowest feasible liability, and work in progress with Arkadii Slinko shows that its index of excess wait does not exceed 1.21. How far is 1.21 from the lowest index of excess wait among all rules with minimal liability?

3. Is Theorem 2 a tight statement? I.e., can we find split-proof rules meeting one or both of Ranking and $\{$for some $n,\ \lambda(n,a) \leq n \cdot a$ for all $a\}$, and different from the Proportional one? I conjecture that such rules exist.

4. In the companion paper Moulin [2004] I discuss the partial transfer of jobs among users, another strategic maneuver related to merging and splitting. Two (or more) agents choose to reallocate their jobs of sizes x_1 and x_2 as x_1' and x_2', with $x_1+x_2 = x_1'+x_2'$, and for instance $x_1 < x_1'$ and $x_2 > x_2'$. If the server schedules job x_1' before job x_2', then agent 1's true job of size x_1 and a chunk of agent 2's true job is processed at that time, and the rest of agent 2's true job is completed when job x_2' is served; if job x_2' comes up before job x_1', both agents must wait until the latter time to complete their true jobs. The transfer of jobs encompasses splitting: if $x_2 = 0$, the above move is equivalent to a split of job x_1 in two; it encompasses merging as well: if $x_2' = 0$, we have the merging of jobs x_1 and x_2 (recall Zero-Consistency in footnote 4).

What scheduling rules are invulnerable to partial transfers among two or more participants? In the quasi-linear model of our companion paper, all efficient rules are vulnerable to transfers involving three or more agents, but pairwise transfer-proofness points to a one-dimensional line of rules borne by two appealing rules, one split-proof and the other merge-proof. These solutions are not built on any proportionality idea, but apply instead the Shapley value to an appropriate cooperative game. They can also be derived from invariance or consistency properties (Maniquet [2003], Chun [2004 a,b]).

I conjecture that the situation is more favourable in the probabilistic scheduling framework. In particular, the Proportional and Uniform rules appear to be transfer-proof, for any number of participants. To understand the broader impact of transfer-proofness is left for future research.

References

[1] Aczél, J. 1996. *Lectures on Functional Equations and Their Applications*. New York: Academic.

[2] Abdulkadiroglu, A. and T. Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems, *Econometrica*, 66, 1998, 689-701.

[3] Banker, R. 1981. Equity consideration in traditional full-cost allocation practices: an axiomatic perspective, *Joint Cost Allocations*, S. Moriarty, ed. Norman: University of Oklahoma.

[4] Bogomolnaia, A. and H. Moulin. A New Solution to the Random Assignment Problem, *Journal of Economic Theory*, 100, 295-328, 2001.

[5] Bogomolnaia, A. and H. Moulin. A Simple Random Assignment Problem with A Unique Solution, *Economic Theory*, 19, 3, 623-636, 2002.

[6] Chun, Y., 2004. A Note on Maniquet's Characterizations of the Shapley Value in Queuing Problems, mimeo.

[7] Chun, Y., 2004. Consistency and Monotonicity in Sequencing Problems, mimeo.

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

[9] Curiel, I., J. Potters, V. Rajendra Prasad, S. Tijs and B. Veltman. Cooperation in One Machine Scheduling, *Methods of Operations Research*, 38, 1993, 113-131.

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

[11] Crés, H. and H. Moulin. Scheduling with Opting Out: Improving upon Random Priority, *Operations Research*, 49, 4, 565-577, 2001.

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

[13] Dolan, R., 1978. Incentive Mechanisms for Priority Queueing Problems, *The Bell Journal of Economics* 9, 421-436.

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

[15] Fishburn, P. Induced binary probabilities and the linear ordering polytope: a status report, *Mathematical Social Sciences*, 23, 1992, 67-80.

[16] Friedman, E. and S. Henderson. Fairness and Efficiency in Processor Sharing Protocols to Minimize Sojourn Times, 2002, mimeo.

[17] Haviv, M. The Aumann-Shapley Price Mechanism for Allocating Congestion Costs, *Operation Research Letters*, 29, 2001, 211-215.

[18] Haviv, M. and Y. Ritov. Externalities, Tangible Externalities, and Queue Disciplines, *Management Science*, 44, 6, 1998, 850-858.

[19] Ju, B.G. Manipulations via Merging and Splitting in claim problems, *Review of Economic Design,* 8, 2003, 205-215.

[20] Ju, B.G. and E. Miyagawa, 2003. Coalitional Manipulations and Generalized Proportional Rules, mimeo, Columbia University.

[21] Kittsteiner, T. and B. Moldovanu, 2003. Auction-Based Queue Disciplines, mimeo, University of Bonn.

[22] Kittsteiner, T. and B. Moldovanu, 2003. Priority Auctions and Queue Disciplines that Depend on Processing Time, mimeo, University of Bonn.

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

[24] Lawler, E., J.K. Lenstra, Kan A. Rinnooy, and D. Shmoys. Sequencing and Scheduling: Algorithms and Complexity, Chapter 9 in Handbooks in *Operations Research* and *Management Science*, Vol. 4, 1993, Elsevier.

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

[26] Mitra, M., 2001. Mechanism Design in Queueing Problems, *Economic Theory,* 17, 277-305.

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

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

[29] Moulin, H., 1985 Egalitarianism and Utilitarianism in Quasi-linear Bargaining, *Econometrica,* 53, 1, 49-67.

[30] Moulin, H., 1987. Equal or Proportional Division of a Surplus, and Other Methods, *International Journal of Game Theory*, 16, 3, 161–186.

[31] Moulin, H., 2004. On Scheduling Fees to Prevent Merging, Splitting and Transferring Jobs, mimeo.

[32] Moulin, H. and Y. Sprumont, 2003. On Demand Responsiveness in Additive Cost Sharing, forthcoming in *Journal of Economic Theory.*

[33] Moulin, H. and R. Stong. Fair Queuing and other Probabilistic Allocation Methods, *Mathematics of Operations Research*, 27, 1, 2002, 1-31.

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

[35] Shapley, L.S., 1971. Core of Convex Games, *International Journal of Game Theory*, 1, 11-26.

[36] Sprumont, Y. Ordinal Cost Sharing, *Journal of Economic Theory*, 81, 1998, 126-162.

[37] Sprumont, Y., 2003. Aumann-Shapley pricing: a reconsideration of the discrete case, mimeo, Universite de Montreal.

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

9 Appendix

9.1 Proof of lemma 1

Statement i). Write $F_0(N,x)$ for the set of feasible profiles of expected waits at (N,x), and $C(N,x)$ for the core of $(N,v(\cdot,x))$ Also use $x \cdot y$ for the coordinatewise multiplication of vectors x,y Fix (N,x) and define $\widetilde{N} = \{i \in N | x_i > 0\}$. Together the three following facts prove the statement:

1.If $N = \widetilde{N}$ then $y \in F_0(N,x) \Longleftrightarrow x \cdot y \in C(N,x)$

2. $y \in F_0(N,x) \Longleftrightarrow \{y_{[\widetilde{N}]} \in F_0(\widetilde{N}, x_{[\widetilde{N}]})$ and $y_i \in [0, \sum_N x_j]\}$

3. $x \cdot y \in C(N,x) \Longleftrightarrow x_{[\widetilde{N}]} \cdot y_{[\widetilde{N}]} \in C(\widetilde{N}, x_{[\widetilde{N}]})$.

Fact 3 requires no proof and fact 2 is clear: once we order jobs randomly in $\widetilde{N}$ to achieve $y_{[\widetilde{N}]}$, where we schedule null jobs does not matter and we can achieve any $y_{[N \diagdown \widetilde{N}]}$ within the announced bounds.

Now fix (N,x) with $x \gg 0$, and for all $\sigma \in \Phi(N)$ let y^σ be the profile of wait under σ namely $y_i^\sigma = \sum_{P(i;\sigma)} x_j$, where $P(i;\sigma) = \{j \in N | \sigma(j) \leq \sigma(i)\}$. Set $P_-(i;\sigma) = P(i;\sigma) \diagdown \{i\}$. Routine computation shows $(x \cdot y^\sigma) = v(P(i;\sigma),x) - v(P_-(i;\sigma),x)$, namely $x \cdot y^\sigma$ is the vector of marginal contributions in the game $(N, v(\cdot,x))$. As $v(\cdot,x)$ is supermodular, a classic result (Shapley [1971]) says that $C(N,x)$ is the convex hull of $x \cdot y^\sigma, \sigma \in \Phi(N)$. On the other hand $F_0(N,x)$ is the convex hull of $y^\sigma, \sigma \in \Phi(N)$. This gives fact 1 because $x \gg 0$.

Statement ii). "Only if" is clear from statement i). "If" follows from the fact that such a vector minimize the sum $\sum_{\widetilde{N}} x_i \cdot y_i + \sum_{N \diagdown \widetilde{N}} y_i$ over $F_0(N,x)$.

9.2 Proof of Proposition 1

Only the " if" statement requires a proof. Let ρ be a rule with the four stated properties, and $\pi(N,x)$ be the associated probability distribution of the last scheduled agent. By Anonymity, it takes the form $\pi(n,x)$, where $n = |N|$ and is symmetric in all variables x_i. Fix n,x and use the simplified notation $\{1,2,..,n-1|x\}$ for the event $\{\sigma(1) < \sigma(2) < .. < \sigma(n-1)|x\}$. A simple partition of this event gives the equality

$$prob\{1,2,..,n-1|x\} = prob\{1,2,..,n|x\} + \sum_{1 \leq i \leq n-1} prob\{1,..,i-1,n,i,..,n-1|x\}$$

The above summation is $\pi_{n-1}(x) \cdot prob\{1,2,..,n-2|x_{[N\setminus n,n-1]}\}$, by Separability and Recursivity. Writing $\pi(n,x)$ simply as $\pi(x)$, we get

$$(1-\pi_n(x))\cdot prob\{1,2,..,n-1|x_{[N\setminus n]}\} = \pi_{n-1}(x)\cdot prob\{1,2,..,n-2|x_{[N\setminus n,n-1]}\} \tag{11}$$

Apply (11) first for $n = 3$, with the notation $\theta(b,a) = \pi_2((a,b))$ for the probability that a job of size a precedes one of size b. We have for all $x = (a,b,c)$, $x \gg 0$, $(1-\pi_3(x)) \cdot \theta(b,a) = \pi_2(x)$. By Positivity, $\theta(b,a) > 0$, so $\pi_2(x) = 0$ would imply $\pi_1(x) = 0$. Exchanging the role of agents 1 and 3, we see that $\pi_2(x) = 0$ is impossible. Thus $\pi_i(x) > 0$ for $i = 1,2,3$. We can now rewrite (11) as

$$\theta(b,a) = \frac{\pi_2}{\pi_1+\pi_2}(a,b,c) \Longleftrightarrow \frac{\pi_1}{\pi_2}(a,b,c) = \frac{1}{\theta(b,a)} - 1$$

namely the ratio $\frac{\pi_1}{\pi_2}$ only depends on a,b. From $\frac{\pi_1}{\pi_2} \cdot \frac{\pi_2}{\pi_3} \cdot \frac{\pi_3}{\pi_1}$, a standard argument gives the existence of three real positive function w_i such that $\frac{\pi_i}{\pi_j}(a,b) = \frac{w_i(a)}{w_j(a)}$ for all i,j in $\{1,2,3\}$. Anonymity gives $w_i = w$ for all i, and we conclude $\pi_i(x) = \frac{w_i(x_i)}{w_1(x_1)+w_2(x_2)+w_3(x_3)}$ for all 3-dimensional profile $x \gg 0$.

Finally we show $\pi_i(x) = \frac{w_i(x_i)}{\sum w_j(x_j)}$ for any positive profile in $\mathbb{R}^n_+$, by an induction argument on n. Assume this holds up to $n-1$. This implies $prob\{1,2,..,n-2|x_{N\setminus n,n-1}\} \neq 0$, therefore (11) becomes $(1-\pi_n(x)) \cdot \pi_{n-1}(x_{N\setminus n}) = \pi_{n-1}(x)$. Now $\pi(x) \gg 0$ and the desired claim follow at once.

9.3 Proof of Lemma 3 statement ii, and Proposition 4

Step 1. We start by a reformulation of the probability that a job of size b precedes one of size a, namely $\theta(a,b)$ given by (3), for a general parametric rule $\{F_a, a > 0\}$ where the cdfs may have atoms. Recall that each F_a lives in the space $\mathcal{H}$ of non-decreasing, right-continuous real-valued functions on $\mathbb{R}_+$ with bounded range. For such a function F we write $\partial F(z) = F(z) - F_-(z)$ for the jump of F at z, where $F_-(z)$ is the left limit of F at z. We denote by F^+ the integral of these jumps, namely the staircase function $F^+(z) = \sum_{t\leq z} \partial F(t)$ (the sum is well defined because F is non-decreasing). Check that $F^0 = F - F^+$ is continuous and non decreasing.

This canonical decomposition, $F = F^0 + F^+$ allows us to define the integral $\int_0^\infty F \cdot dG$, for any two $F, G \in \mathcal{H}$. If F, G are both continuous, this is the familiar Stieljes integral (e.g., Hardy et al [1938]). For arbitrary F, G in $\mathcal{H}$ we write $\mathcal{J}(F), \mathcal{J}(G)$ for the set of their jumps, namely $z \in \mathcal{J}(F) \Leftrightarrow \partial F(z) > 0$; then we define $\int F^+ \cdot dG^0 = \sum_{\mathcal{J}(F)} \partial F(z) \cdot (G(\infty) - G(z)) = (F^+ \cdot G)(\infty) - \sum_{\mathcal{J}(F)} \partial F(z) \cdot G(z)$, and $\int F \cdot dG^+ = \sum_{\mathcal{J}(G)} F(z) \cdot \partial G(z)$. these combine into the following definition

$$\int_0^\infty F \cdot dG = \int_0^\infty F^0 \cdot dG^0 + (F^+ \cdot G^0)(\infty) - \sum_{\mathcal{J}(F)} \partial F(z) \cdot G^0(z) + \sum_{\mathcal{J}(G)} F(z) \cdot \partial G(z).$$

From this it is straightforward to deduce the integration by parts formula:

$$\int_0^\infty F \cdot dG + \int_0^\infty G \cdot dF = (F \cdot G)(\infty) - (F \cdot G)(0) + \sum_{\mathcal{J}(F) \cap \mathcal{J}(G)} \partial F(z) \cdot \partial G(z). \tag{12}$$

Now we take two random variables Z_a, Z_b with cdfs F_a, F_b, and we compute

$$prob\{Z_a = Z_b\} = \sum_{\mathcal{J}(a) \cap \mathcal{J}(b)} \partial F_a(z) \cdot \partial F_b(z)$$

indeed the equality only occurs when both draws are in the atomic part of the two cdfs. Similarly the probability of $\{Z_b \leq Z_a\}$ is

$$\int_0^\infty F_b^0 \cdot dF_a^0 + (F_b^+ \cdot F_a^0)(\infty) - \sum_{\mathcal{J}(b)} \partial F_b(z) \cdot F_a^0(z) + \sum_{\mathcal{J}(a)} F_b(z) \cdot \partial F_a(z) = \int_0^\infty F_b \cdot dF_a$$

where the last term is the probability of $\{Z_b \leq Z_a\} \cap \{Z_a \in \mathcal{J}(a)\}$ and the sum of the two middle terms is the probability of $\{Z_b \leq Z_a\} \cap \{Z_b \in \mathcal{J}(b)\}$. Comparing with equation (3), and using integration by parts (12) we conclude

$$\begin{aligned} \theta(a, b) &= \int_0^\infty F_b \cdot dF_a - \frac{1}{2} \sum_{\mathcal{J}(a) \cap \mathcal{J}(b)} \partial F_a(z) \cdot \partial F_b(z) = \\ &= \int_0^\infty (1 - F_a) \cdot dF_b + \frac{1}{2} \sum_{\mathcal{J}(a) \cap \mathcal{J}(b)} \partial F_a(z) \cdot \partial F_b(z) \end{aligned} \tag{13}$$

Step 2: statement ii) in Lemma 3. Let $\{F_a, a > 0\}$ be as in the statement. We develop $\theta(a,b)$ with the help of (13) and the definition of the integral

$$\theta(a,b) = \int_0^\infty (1-F_a)\cdot dF_b^0 + \sum_{\mathcal{J}(b)} (1-F_a)(z)\cdot \partial F_b(z) + \frac{1}{2}\sum_{\mathcal{J}(a)\cap\mathcal{J}(b)} \partial F_a(z)\cdot\partial F_b(z)$$

The first integral term is non decreasing in a as $F_a(z)$ is non increasing in a. The second term is also non decreasing, but the variation of $\partial F_a(z)$ in a is ambiguous. However, the terms corresponding to any $z \in \mathcal{J}(a)\cap\mathcal{J}(b)$ in the second and third sum are

$$(1 - F_a(z))\cdot\partial F_b(z) + \partial F_a(z)\cdot\partial F_b(z) = 1 - \frac{1}{2}(F_a(z) + F_b^-(z))\cdot\partial F_b(z)$$

and the desired monotonicity follows.

Step 3: Proposition 4. We must show inequality (9) when $F_a \cdot F_b \leq F_{a+b}$ holds for all $a, b > 0$. We must prove inequality (9), which takes the following form for parametric rules (as already discussed just before Proposition 4).

$$\overline{a} = \sum_1^m a_j \Longrightarrow prob\{Z_{\overline{a}} \leq Z_b\} \geq prob\{\max_{1,..,m} Z_{a_j} \leq Z_b\} \tag{14}$$

Denote by $\widetilde{Z}_k$ the k-th order statistics of the independent variables Z_{a_j}, thus $\widetilde{Z}_1 = \min_{1,..,m} Z_{a_j}$ and $\widetilde{Z}_m = \max_{1,..,m} Z_{a_j}$. Compute

$$\begin{aligned} prob\{\max_{1,..,m} Z_{a_j} &\leq Z_b\} = prob\{\widetilde{Z}_m < Z_b\} + \frac{1}{2}prob\{\widetilde{Z}_{m-1} < \widetilde{Z}_m = Z_b\} \\ +\frac{1}{3}prob\{\widetilde{Z}_{m-2} &< \widetilde{Z}_{m-1} = \widetilde{Z}_m = Z_b\} + \cdot\cdot \leq prob\{\widetilde{Z}_m < Z_b\} + \frac{1}{2}prob\{\widetilde{Z}_m = Z_b\} = \end{aligned}$$

$$= \int_0^\infty F_{a_1}\cdot ..\cdot F_{a_m}\cdot dF_b - \tfrac{1}{2}\textstyle\sum_{\mathcal{J}^*\cap\mathcal{J}(b)} \partial(F_{a_1}\cdot ..\cdot F_{a_m})(z)\cdot\partial F_b(z)$$

where $\mathcal{J}^* = \cup_{1,..,m}\mathcal{J}(a_j)$. Mimicking the argument in Step 2 it is easy to show that $T(G) = \int G\cdot dF - \frac{1}{2}\sum \partial G\cdot\partial F$ is monotonic in the sense that $\{G_1(z) \leq G_2(z) \text{ for all } z\} \Longrightarrow T(G_1) \leq T(G_2)$. Applying this to $G_1 = F_{a_1}\cdot ..\cdot F_{a_m}$ and $G_2 = F_{\overline{a}}$ gives (14) and completes the proof.

9.4 Proof of Lemma 5

We fix a recursive rule ρ as in the statement of lemma 5, and we must prove inequality (8) under the premises of Definition 9. In view of equation (1) it is enough to prove $prob\{\sigma(j) < \sigma(i_*)|x\} \leq prob\{\sigma(j) < \max_{i_* \cup T} \sigma(i)|x_*\}$ which we rephrase as

$$prob\{\sigma(i_*) < \sigma(j)|x\} \geq prob\{\max_{i_* \cup T} \sigma(i) < \sigma(j)|x_*\} \tag{15}$$

. We use an induction argument on the number q of agents in $N \cup T$. For $q = 3$, we must check for all agents $1, 2, 3$, all $a, b, x_3 > 0$, and $x_1 = a + b$

$$prob\{\sigma(1) < \sigma(3)|(x_1, x_3)\} \geq prob\{\sigma(1), \sigma(2) < \sigma(3)|(a, b, x_3)\}$$

which amounts to $\pi_3(x_1, x_3) \geq \pi_3(a, b, x_3)$ and follows by assumption. For the induction step, we use Recursivity to develop both sides of inequality (15), which is then equivalent to the following

$\pi_j(x)+\sum_{N\diagdown i_*,j} \pi_k(x) \cdot prob\{\sigma(i_*) < \sigma(j)|x_{[N\diagdown k]}\} \geq \pi_j(x_*)+\sum_{N\diagdown i_*,j} \pi_k(x_*) \cdot prob\{\max_{i_* \cup T} \sigma(i) < \sigma(j)|(x_*)_{[N \cup T\diagdown k]}\}$.

The subadditivity assumption on π ensures $\pi_j(x_*) \geq \pi_j(x)$ and $\pi_k(x_*) \geq \pi_k(x)$; together with the the inductive assumption this proves (15) and the Lemma.

9.5 Lemma 9 and proof of Lemma 8 statement i

Lemma 9.
Let ρ be a split-proof scheduling rule, and let $\theta^{i,j}(a, b) = prob\{\sigma(j) < \sigma(i)|(a, b)\}$ be the probability that a job of size b by j precedes one of size a by i in the two person problem. Then $\theta^{i,j} = \theta$ is independent of i, j. Moreover $a \to \theta(a, c - a)$ is subadditive in a. Finally $\sup_b \theta(a, b) \cdot b \geq a$.

Proof
That $\theta^{i,j}$ does not depend on i, j was established in the proof of statement "only if"in Lemma 4 (note that Separability of ρ is not needed for the argument). Next we define for all a, c such that $0 \leq a \leq c$, $f(a, c) = \theta(a, c - a)$. Fix $a_1, a_2, a_3 > 0$ and consider the split of agent 2 in problem $\{1, 2\}, (a_1, a_2 + a_3)$ to agents $2, 3$ in $\{1, 2, 3\}, (a_1, a_2, a_3)$. Split-proofness implies

$$f(a_1, a_1 + a_2 + a_3) = \theta(a_1, a_2 + a_3) \geq prob\{1 \text{ is last in } (a_1, a_2, a_3)\} \quad (16)$$

A similar argument gives $f(a_i, a_1 + a_2 + a_3) \geq prob\{i \text{ is last in } (a_1, a_2, a_3)\}$. Summing up gives $\sum_{1,2,3} f(a_i, a_1 + a_2 + a_3) \geq 1$, implying

$$f(a_1, a_1{+}a_2{+}a_3){+}f(a_2, a_1{+}a_2{+}a_3) \geq 1{-}f(a_3, a_1{+}a_2{+}a_3) = f(a_3, a_1{+}a_2{+}a_3)$$

where the equality comes from $\theta(a, b) + \theta(b, a) = 1$.

To prove the last statement, fix $a > 0$ and $m \in \mathbb{N}$, and note $f(m \cdot a, 2m \cdot a) = \frac{1}{2}$. By superadditivity of f, $f(a, 2m \cdot a) \geq \frac{1}{m} f(m \cdot a, 2m \cdot a) \Leftrightarrow \theta(a, (2m-1) \cdot a) \geq \frac{1}{2m}$, from which $\sup_b \theta(a, b) \cdot b \geq a$ follows at once.

Statement i in Lemma 8

If the rule ρ is separable, its liability $\lambda(n, a) = \sup_{x_{-i}} y_i(a, x_{-i})$ is given by equation (10). Now $\sup_b \theta(a, b) \cdot b \geq a$ implies $\lambda(n, a) \geq n \cdot a$.

9.6 Theorems 1 and 2

Throughout the proof, we use the notations of lemma 9, and the properties of the functions θ f discussed there.

Statement i

If ρ is split-invariant, inequality (16) in the above proof is in fact an equality, therefore $\sum_{1,2,3} f(a_i, a_1 + a_2 + a_3) = 1$. This implies that $f(a, c)$ is linear in a (Cauchy's theorem applies because $f \geq 0$), and from $f(a, 2a) = \frac{1}{2}$ (or $f(a, a) = 1$) we get $f(a, c) = \frac{a}{c} \Leftrightarrow \theta(a, b) = \frac{a}{a+b}$. Now if ρ is separable (Theorem 1), its method μ given by (2) is Proportional. Next for Theorem 2 assume ρ is recursive (not necessarily separable). In the split of 2 in problem $\{1, 2\}, (x_1, \sum_{N \diagdown 1} x_i)$ to $\{2, 3, .., n\}$ in problem N, x, Split-invariance gives $\theta(x_1, \sum_{N \diagdown 1} x_i) = \pi_1(x)$, where π is the distribution of the agent served last (Definition 6). We conclude that π is proportional, and by Recursivity that ρ is the Proportional rule.

Statement ii

Fix a split-proof rule ρ such that $\theta(a, b) \leq \frac{a}{b}$ for all $a, b > 0$. We claim that $\theta(a, b) = \frac{a}{a+b}$. The assumption writes $f(a, c) \leq \frac{a}{c-a}$ for all $0 < a < c$. Apply k times the subadditivity of f in its first variable:

$$f(a, c) \leq 2f(\frac{a}{2}, c) \leq .. \leq 2^k f(\frac{a}{2^k}, c) \leq \frac{a}{c - \frac{a}{2^k}}$$

Thus $f(a,c) \leq \frac{a}{c}$. By subadditivity again

$$1 = f(c,c) \leq f(a,c) + f(c-a,c) \leq \frac{a}{c} + \frac{c-a}{c} = 1$$

hence the claim.

To statement $ii)$ in Theorem 1. Fix a split-proof and separable rule ρ such that for some n we have $\lambda(n,a) \leq n \cdot a$, for all a. At the $n-$profile $x = (a,b,b,..,b)$ this implies $y_1(n,x) = a + (n-1) \cdot \theta(a,b) \cdot b \leq n \cdot a$, hence $\theta(a,b) \leq \frac{a}{b}$, and the above observation gives $\theta(a,b) = \frac{a}{a+b}$, so the method derived from ρ is Proportional. Next observe by Lemma 6 that Ranking.implies $\lambda(2,a) \leq 2 \cdot a$ for all a, which completes the proof of Theorem 1.

From the assumptions in statement $ii)$ of Theorem 2, we get as above $\theta(a,b) = \frac{a}{a+b}$. Consider next the split of 2 in problem $\{1,2\}, (x_1, \sum_{N \diagdown 1} x_i)$ to $\{2,3,..,n\}$ in problem N, x. Split-proofness gives $\pi_1(x) \leq \theta(x_1, \sum_{N \diagdown 1} x_i) = \frac{x_1}{\sum_N x_i}$. Repeat the argument for all i: as $\pi(x)$ is a probability distribution, we conclude that it is the Proportional one, and that ρ is the Proportional rule by Recursivity.