

# Fair Queuing and other Probabilistic Allocation Methods

Hervé Moulin and Richard Stong

Rice University

April 2000, Revised March 2001

## Abstract

A server processes one job per unit of time and randomly schedules the jobs requested by a given set of users; each user may request a different number of jobs.

Fair queuing (Shenker [1989]) schedules jobs in successive round-robin fashion, where each agents receives one unit in each round until his demand is met and the ordering is random in each round. Fair queuing\*, the reverse scheduling of fair queuing, serves first (with uniform probability) one of the users with the largest remaining demand.

We characterize Fair Queuing\* by the combination of Lower Composition –LC– (the scheduling sequence is history-independent), Demand Monotonicity –DM– (increasing my demand cannot result in increased delay) and two equity axioms, Equal Treatment Ex Ante –ETEA– (two identical demands give the same probability distribution of service) and Equal Treatment Ex Post –ETEP– (two identical demands must be served in alternating fashion). The set of dual axioms, (in which ETEA and ETEP are unchanged) characterizes Fair Queuing.

We also characterize the rich family of methods satisfying LC, DM and the familiar Consistency –CSY– axiom. They work by fixing a standard of comparison (preordering) between a demand of  $x_i$  units by agent  $i$  and one of  $x_j$  units by agent  $j$ . The first job scheduled is drawn from of the agents whose demand has the highest standard.

*Keywords:* probabilistic allocation, scheduling, fair queuing, consistency, lower and upper composition

Hervé Moulin, Department of Economics, MS 22, Rice University, P.O. Box 1892, Houston, Texas 77251-1892, USA, moulin@rice.edu, <http://www.ruf.rice.edu/~econ/papers/index.html>

Richard Stong, Department of Mathematics, MS 136, Rice University, P.O. Box 1892, Houston, Texas 77251-1892, USA, stong@rice.edu

JEL: D63, D70

*Acknowledgements:* We are grateful to seminar participants at Stanford University, Southern Methodist University, Université d’Aix-Marseille and Rice University, as well as to an anonymous referee. Special thanks to Rakesh Vohra who suggested the result described in Remark 3.

This research is supported by the NSF under grant SES0096230.

# Fair Queuing and other Probabilistic Allocation Methods

## 1 Scheduling/Rationing indivisible units

Scheduling a list of homogeneous indivisible tasks is always unfair ex post, but ex ante fairness can be achieved by lotteries, i.e., by choosing the schedule at random.

A server faces demands of various sizes from various users. Assume that user  $i$ 's demand is of size  $x_i$ , and that our server performs one task per unit of time; in other words it can meet one unit of demand from any agent per unit of time. For instance, messages sent on the internet are broken into a series of elementary packets, and a switch processes packets sequentially (see Demers et al. [1990]). The scheduling problem is to choose the order in which the elementary tasks will be performed; here again randomization allows ex ante fairness.

In our simple model, the list of demands  $x_i$ , one per agent, is deterministic and fixed once and for all: we do not allow for the random arrival of new demands at successive dates. For instance the model cannot describe the familiar FIFO method.

Two simple scheduling methods are discussed in the queuing literature. The *proportional* method amounts to treat equally each unit of claim: for each agent  $i$  we throw  $x_i$  balls labeled  $i$  in an urn, and draw the balls successively (without replacement and with uniform probability among the remaining balls). Equivalently, the  $t$ -th unit goes to agent  $i$  with a probability proportional to his remaining (unsatisfied) demand.

The proportional method is a simple and natural interpretation of fairness in our problem. Yet it has been criticized in the scheduling problem because it fails to protect an agent with a small demand. This is because an agent's expected waiting time goes to infinity if other agents' demands do the same. The *fair queuing* method<sup>1</sup> solves this problem by allocating one unit per agent, irrespective of the size of individual demands, in successive round-robin fashion. In each round the active agents (whose demands is not yet fully met) are randomly ordered (with uniform probability) and served one unit (one job) in that order. Fair queuing also stands out for its superior incentives properties: inflating one's demand artificially to increase one's chance of early service does not pay – as it would under the proportional method –.

An equivalent interpretation of our model is the problem of rationing an overdemand commodity, the object of a substantial amount of axiomatic analysis in the microeconomic literature on distributive justice<sup>2</sup>.

---

<sup>1</sup>Originally proposed by Shenker [1989] in the context of networks; see Demers et al. [1990] for references to that literature.

<sup>2</sup>References on that literature are given in the next section.

In the rationing interpretation, agent  $i$  demands  $x_i$  units of the commodity, and the available resources  $t$  fall short of total demand. A *probabilistic rationing method* allocates the resources possibly with the help of a lottery, making sure that no agent receives more than his demand.

Admission to colleges, the award of immigration visas, the allocation of workers between the divisions of a firm are all instances of rationing where the overdemanded resources come in indivisible units. Lotteries are the simplest and oldest mechanism to restore (ex ante) fairness despite the indivisible character of the resources.<sup>3</sup>

When  $t$  increases from  $t = 1$  to  $t = \sum x_i$ , a plausible requirement is that the (random) share of a given agent never decreases. Granting this property<sup>4</sup>, the rationing of the profile of demands  $x_i$  is described by a (random) sequence of agents  $\{i_1, i_2, \dots\}$  where each agent  $i$  appears exactly  $x_i$  times (thus the length of the sequence is  $\sum x_i$ ):  $i_1$  is the agent receiving the unit if  $t = 1$ , agent  $i_2$  receives the second unit and so on. If the demand  $x_i$  represents the number of “jobs” requested by agent  $i$ , such a sequence is precisely a schedule of these jobs.

Keeping in mind both interpretations of the formal model, rationing and scheduling, is helpful throughout the paper to interpret the normative requirements (axioms) as well as the results.

A very general equity requirement is Equal Treatment of Equals:  $x_i = x_j \Rightarrow y_i = y_j$ , namely unequal shares can only be justified by unequal demands. If ETE fails, the allocation method embeds exogenous differences (unequal rights) among agents<sup>5</sup>. In real life the equity requirement is often compelling (as in “one man, one vote”) but inequitable allocation methods are important as well: differences in priority status between various users of the internet and other networks is an important policy tool to control congestion. From the standpoint of axiomatic analysis, we want accordingly to study equitable rationing methods as well as inequitable ones.

In our probabilistic rationing model, agent  $i$ 's demand  $x_i$  and the resources  $t$  are (deterministic) non negative integers; the share  $Y_i$  assigned to agent  $i$  by a rationing method is an integer valued random variable such that  $0 \leq Y_i \leq x_i$  and  $\sum_i Y_i = t$ . The “equal treatment of equals” principle takes two interestingly different forms. In the ex ante form (ETEA for short), we require  $x_i = x_j \Rightarrow Y_i \sim Y_j$  in the sense that these two variables have identical probability distributions. Both rationing methods, *proportional* and *fair queuing* meet ETEA.

In the ex post form (ETEP) we require  $x_i = x_j \Rightarrow |Y_i - Y_j| \leq 1$  namely the realization of the two main variables can differ by one at most. This is the

---

<sup>3</sup>See Hofstee [1990] for a college admission example, and Elster [1992] chapter 2, Young [1994] chapter 2 for more examples, including demobilization and organ transplants.

<sup>4</sup>The precise link between probabilistic rationing and scheduling is discussed in Section 5: see in particular Remark 3.

<sup>5</sup>Stronger equity requirements include Ranking:  $x_i \leq x_j \Rightarrow y_i \leq y_j$ , and Anonymity, expressing that the mapping  $x \rightarrow y$  is symmetric. In our model, these strong equity properties are a consequence of ETE, in combination with the other axioms.

smallest feasible difference in (ex post) shares. For instance, if  $x_i = x_j$  for all  $i, j$  and  $t$  is a multiple of  $n$ , the number of agents, ETEP forces deterministic equal shares. *Fair queuing* meets ETEP, but *proportional* does not: for instance, if  $t \leq x_i$  for all  $i$ , this method gives everything to a single agent with positive probability.

Our results rely on three powerful invariance axioms familiar in the rationing literature. Consistency (CSY) says that the allocation of  $t$  units among agents  $\{1, 2, \dots, n\}$  can be made in two stages: first we serve (a random share to) agent 1, then we share the remaining units between agents  $\{2, \dots, n\}$ . In the scheduling interpretation, CSY simply says that dropping an agent does not alter the scheduling of the remaining agents.

The second axiom is Lower Composition (LC), stating that the  $t$  units can be handed out one at a time, provided each time we decrease by one the demand of the recipient. Equivalently LC says that the scheduling process is “history-independent”: at any point in time, the size of the remaining demands is all that matters to the scheduling of the remaining jobs – the particular realization of the earlier sequence of jobs is irrelevant. The third axiom, Upper Composition (UC), is the dual of LC: it says that the  $t^*$  units of deficit ( $t^* = \sum_i x_i - t$ ) can be handed down one at a time, provided we decrease accordingly the demand of the recipient. Equivalently, the scheduling process is “future-blind”: for any given  $t$ , the total accumulation of jobs up to time  $t$  (i.e., how many jobs are assigned to each agent  $i$  at  $t$ ) is all we need to compute the random scheduling of the first  $t$  jobs.

Our first main result, Theorem 1, shows that the three axioms Lower and Upper Composition, and Equal Treatment Ex Post, drive a wedge between the three basic methods. The proportional one meets LC and UC but fails ETEP (see above). Fair Queuing meets ETEP and UC but fails LC. The successive rounds allocating one unit per agent still active, are not history-independent: one must keep track of who has been served so far in the current round. Thus LC fails. On the other hand UC holds because fair queuing allocates successive units of deficit uniformly among the agents with the largest remaining demand. The dual Fair Queuing\* method meets ETEP and LC but not UC.

Theorem 1 offers compact characterizations of our three methods in which equity is assumed but Consistency plays no role (although this property is satisfied by all three). The other two main results, Theorem 2 and Theorem 3, drop the requirement of equity and essentially characterize the rich families of methods satisfying any two out of the three properties CSY, LC and UC.

In order to describe these rich and somewhat complex families, we need one more axiomatic ingredient, namely the two mild properties called Demand Monotonicity (DM) and its dual Demand Monotonicity\* (DM\*). DM says that an increase in agent  $i$ 's demand  $x_i$ , leaving  $t$  and the other demands unchanged, cannot deteriorate agent  $i$ 's (random) share, in the sense of stochastic dominance. The dual axiom DM\* requires that agent  $i$ 's (random) share of deficit does not diminish as his demand increases while the deficit  $t^*$  and the other demands are unchanged.

Theorem 2 characterizes the set of methods meeting CSY, LC and DM; or,

dually, the methods meeting CSY, UC and DM\*. In order to describe this rich family, it is convenient to start with the simple subset of its deterministic methods, i.e., those methods that do not use any lottery. In this case the key concept is that of *standard of gains*<sup>6</sup>, namely an ordering of all pairs  $(i, x_i)$  made of one agent and one positive claim. The relation  $(i, x_i) \succ (j, x_j)$  reads “a claim of  $x_i$  by agent  $i$  has precedence over a claim of  $x_j$  by agent  $j$ ”. A standard of gain is a linear ordering (no indifferences) of the pairs  $(i, x_i)$  that is increasing in claims:  $(i, x_i) \prec (i, x_i + 1)$ . The rationing method associated with a standard of gains allocates the first unit to the agent  $i$  such that  $(i, x_i)$  is highest for the standard, and reduces agent  $i$ 's claim to  $x_i - 1$ ; the second unit goes similarly to agent  $j$  (possibly the same) such that  $(j, x_j)$  is highest – given agent  $i$ 's reduced claim – and reduces that agent's claim by one; and so on.

Proposition 1 says that in the deterministic case, the trio CSY, LC and DM characterizes the family of standard of gains methods. The dual Proposition 1\* says that CSY, UC and DM\* characterizes the set of standard of losses methods, where the above algorithm is used to allocate the successive units of deficit.

When lotteries are allowed, a standard of gains (or losses) allows indifferences, hence it is simply a preordering (transitive, complete) over all pairs  $(i, x_i)$ , non decreasing in  $x_i$ . Now a typical rationing method associated with a standard of gains works as follows. Given the profile of claims  $x_i$ , we identify all pairs  $(i, x_i)$  maximizing the standard. If there is only one pair  $(i, x_i)$ , agent  $i$  gets the first unit and his claim drops by one; if the pairs with highest standard are, say,  $(1, x_1)$ ,  $(2, x_2)$  and  $(3, x_3)$ , we seek the lowest claims  $x'_1$ ,  $x'_2$  and  $x'_3$  such that  $(1, x'_1)$ ,  $(2, x'_2)$  and  $(3, x'_3)$  are still in the same indifference class say  $(1, x'_1)$ ,  $(2, x'_2)$  and  $(3, x'_3)$ . Then we use the proportional method to allocate the first  $t = 3 + x_1 - x'_1 + x_2 - x'_2 + x_3 - x'_3$  units among agents 1, 2, 3 with initial claims  $(x_1 - x'_1 + 1, x_2 - x'_2 + 1, x_3 - x'_3 + 1)$ . After  $t$  units have been handed out in this way, the remaining claims are  $x'_1 - 1, x'_2 - 1, x'_3 - 1$ , and we repeat the algorithm to allocate the next units.

If the standard of gains makes all pairs  $(i, x_i)$  indifferent, the proportional method obtains. The standard of gains  $(i, x_i) \succeq (j, x_j)$  iff  $x_i \geq x_j$  yields the fair queuing\* method, where the first unit goes with equal probability to one of the agents with the largest demand.

Theorem 2 characterizes the family of standard of gains methods<sup>7</sup> by the combination of the axioms CSY, LC and DM. Dually, the trio CSY, UC and DM\* characterizes the set of *standard of losses* methods, where the above algorithm is used to distribute the successive units of deficit. For instance fair queuing corresponds to the standard of losses  $(i, x_i) \succeq (j, x_j)$  iff  $x_i \geq x_j$ .

Our last result, Theorem 3, bears on the family of rationing methods meeting Lower and Upper Composition. This set contains only one equitable method, namely proportional (Theorem 1 in Moulin [1999a]). Even under the additional requirements DM and DM\*, the family in question is quite complicated and does not appear to have a simple representation, except in the deterministic

<sup>6</sup>Originally introduced by Young [1994]; see also Kaminski [2000].

<sup>7</sup>The precise definition of these methods is only a little bit more involved than the above discussion suggests: see Section 7.

case: Proposition 3. However, if we restrict attention to those methods in which every agent with a positive demand receives more than nothing (at least one unit with a positive probability)<sup>8</sup>, the family defined by LC+UC has an elegant representation. Every such method is entirely determined by a set of  $n(n-1)/2$  numerical weights ( $n$  is the number of agents), via a familiar combinatorial formula.

## 2 Related literature

In the classical rationing problem<sup>9</sup>, demands and shares are both deterministic and divisible. Agent  $i$ 's demand  $x_i$  of the divisible commodity is a real number and so is  $t$ , the amount available for distribution. A rationing method assigns a real number  $y_i$  to agent  $i$ , his share, where  $0 \leq y_i \leq x_i$  and  $\sum_i y_i = t$ .

The three methods called Proportional, Uniform Gains (or Constrained Equal Award), and Uniform Losses (or Constrained Equal losses) stand out in the axiomatic discussion. The proportional shares are  $y_i = \frac{x_i}{\sum_j x_j}$ . The uniform gains shares are  $y_i = \min\{x_i, \lambda\}$ , where  $\lambda$  is the solution of  $\sum_j \min\{x_j, \lambda\} = t$ . The uniform losses shares are  $y_i = \max\{x_i - \mu, 0\}$ , where  $\mu$  solves  $\sum_j \max\{x_j - \mu, 0\} = t$ .

In our model with indivisible units and probabilistic shares, the expected shares generated by the Proportional method are clearly  $y_i = EY_i = t \cdot \frac{x_i}{\sum_j x_j}$ . It is just as easy to check that the expected shares of Fair Queuing (resp. Fair Queuing\*) are precisely the uniform gains (resp. uniform losses) shares. Furthermore, our three invariance axioms –CSY, LC and UC – mimic the axioms with the same name in the classical model.

Due to the discrete nature of the model, our results are sharper and simpler than in the classical model. For instance our compact characterization of Fair Queuing and Fair Queuing\* (statements  $b$  and  $c$  in Theorem 1) has no counterpart in the classical model.

An important result due to H. P. Young illustrates clearly the differences between the two models. Theorem 1 in Young [1988] characterizes a family of “equal sacrifice” methods by the combination of CSY, UC, a very strong equity axiom (Strict Ranking:  $x_i < x_j \Rightarrow y_i < y_j$ ) and a monotonicity axiom (Strict Resource Monotonicity:  $t < t' \Rightarrow x_i < x'_i$ ). The latter axiom has no counterpart in our model with indivisible goods. Nor are the equal sacrifice methods discussed by Young related in any clear way to the standard of losses methods characterized in our Theorem 2 by the combination of CSY, UC and DM\*. For instance, Uniform Losses (the analog of Fair Queuing\*) is ruled out by Strict Ranking – and by Strict Resource Monotonicity as well.

Closer to home, Moulin [2000] characterizes the impact of the trio CSY + LC + UC, both in the classical model and in the model with indivisible goods

<sup>8</sup>The proportional is one such method, fair queuing is not.

<sup>9</sup>The literature starts with Banker [1981], O'Neill [1982] and Aumann Maschler [1985]. Two surveys are Thomson [1995] and Moulin [1999c].

and deterministic shares. In the latter, this trio is met only by the priority methods (see Remark 1 in Section 4). In the former, the set of such methods is large, richer than in the current probabilistic model: for instance there are exactly three equitable methods meeting CSY, LC, UC, and Scale Invariance, a natural homogeneity property with no counterpart in our discrete model. These three methods are precisely the proportional, Uniform Gains and Uniform Losses methods (Corollary of Theorem 2 in Moulin [2000]), a result easily compared to our Theorem 1.

The current rationing model with indivisible units and probabilistic shares is introduced in Moulin [1999a]. That paper provides a characterization of the Proportional method reproduced below as statement *a* in Theorem 1 as well as a characterization of the family of methods meeting CSY, LC and UC. This result is described in the concluding Section 10.

Sasaki [1996] introduces a model closely related to ours, where homogeneous indivisible units are allocated randomly. In Sasaki's model agents have single-peaked preferences over their share, and the sum of individual peaks  $x_i$  can be larger or smaller than the available resources  $t$  (the commodity can be over-demanded or oversupplied). He characterizes the (analog of the ) Fair Queuing method by the combination of equity, efficiency, and incentive compatibility (strategyproofness): neither of the latter concepts applies to our preference-free context.

#### *Organization of the paper*

Section 3 defines the deterministic model. Section 4 introduces the deterministic standard of gains/losses methods and states the deterministic version of Theorem 2: Propositions 1 and 1\*. The probabilistic rationing and scheduling models are introduced in Section 5. The proportional, fair queuing and fair queuing\* methods are characterized in Section 6: Theorem 1. Section 7 describes the probabilistic standard of gains/losses methods; they are characterized in Section 8: Theorem 2. The final Section 9 discusses the methods meeting LC and UC: Theorem 3. All proofs are gathered in the Appendix.

### **3 Deterministic model**

The finite set of agents  $N$  is fixed throughout the paper. We denote by  $\mathbb{N}$  the set of integers,  $\mathbb{N} = \{0, 1, 2, \dots\}$ . Agent  $i$ 's demand  $x_i$  and the resources  $t$  are in  $\mathbb{N}$ . The  $i$ -th unit vector of  $\mathbb{N}^N$  is denoted  $e_i$ . The dimension of a vector  $x$  in  $\mathbb{N}^N$  is the number of its non zero coordinates. We denote by  $x$  a profile of demands,  $x \in \mathbb{N}^N$ , and write  $x_N = \sum_i x_i$ . A rationing problem is a pair  $(t, x)$ , such that  $t \leq x_N$ . A rationing method  $r$  associates to every problem  $(t, x)$  a vector of shares  $y = r(t, x)$  in  $\mathbb{N}^N$  and such that  $0 \leq y_i \leq x_i$  for all  $i$  and  $y_N = t$ .

The dual  $r^*$  of a method  $r$  allocates the  $t$  units of resource in the same way as  $r$  allocates  $x_N - t$  units of deficit



$$r^*(t, x) = x - r(x_N - t, x) \quad (1)$$

All methods discussed below are *resource monotonic*, namely  $t \leq t' \implies r(t, x) \leq r(t', x)$ . For instance, this property follows from either one of Lower or Upper Composition defined below.

In the scheduling interpretation, a resource monotonic rationing method associates to every demand profile  $x$  a sequence of length  $x_N$  in  $N$ , denoted  $r(x) = \{i_1, i_2, \dots, i_{x_N}\}$ , such that each agent  $i$  appears exactly  $x_i$  times. The  $i_t$  unit vector  $e_{i_t}$  is the increment  $r(t, x) - r(t-1, x)$ . We shall use both formulations of a method  $r$  – rationing and scheduling – indifferently. For instance in the scheduling format, the dual sequence  $r^*(x)$  is simply the reverse sequence  $j_t = i_{x_N - t + 1}$ .

The next two properties respectively state that an agent's share does not decrease when his demand rises, *ceteris paribus* (axiom DM), and that his share of deficit does not decrease when his demand rises and *we keep the deficit constant* (axiom DM\*).

Demand Monotonicity (DM):  $r_i(t, x) \leq r_i(t, x + e_i)$

Demand Monotonicity\* (DM\*):  $r_i(t+1, x + e_i) \leq r_i(t, x) + 1$

As suggested by our notation, DM\* is the dual of DM: a method  $r$  meets DM\* if and only if  $r^*$  meets DM.

In the deterministic model, we introduce a single equity axiom initially proposed by Balinski and Young [1982] in the related apportionment problem:

Equal Treatment Ex Post (ETEP):  $x_i = x_j \implies |r_i(t, x) - r_j(t, x)| \leq 1$

In the scheduling context, ETEP requires that two agents with equal demands are served alternately.

The next three axioms are the backbone of our analysis. We use the notation  $(x \mid^i y_i)$  for the following vector  $z$  in  $\mathbb{N}^N$ :  $z_i = y_i, z_j = x_j$  for  $j \neq i$ .

Consistency (CSY)<sup>10</sup>  $r_i(t, x) = r_i(t - r_j(t, x), (x \mid^j 0))$  for all  $i \neq j$ .

The rationing interpretation of CSY is well-known: instead of sharing  $t$  among all agents in  $N$ , we may give her share to agent  $j$ , then cancel  $j$ 's demand and divide the remaining  $(t - y_j)$  units among  $N \setminus j$ .

In the scheduling format, CSY takes a very simple form: the sequence  $r((x \mid^j 0))$  obtains from  $r(x)$  by deleting agent  $j$ 's occurrences. By repeated application, we get the following equivalent formulation of the CSY property: for any pair  $(i, j)$ , the  $(i, j)$  subsequence of  $r(x)$  (obtained from  $r(x)$  by deleting all agents other than  $i, j$ ) is independent of  $x_k$  for all  $k \neq i, j$ .

Lower Composition (LC):  $r(t, x) = r(1, x) + r(t-1, x - r(1, x))$

<sup>10</sup>The standard formulation of CSY uses a model with variable population  $N$  and relates the choices in  $N$ - problems and in  $N \setminus i$ - problems (see e.g., Thomson [1995], Moulin [1999c]). In our model an agent  $i$  with null demand receives  $y_i = 0$ , hence a problem  $(N, t, x)$  where  $x_i = 0$  can be identified with the problem  $(N \setminus i, t, x_{-i})$ . Thus our formulation of CSY coincides with the standard one.

By repeated applications of LC, we obtain the following equivalent property. The allocation of  $t$  units at the demand profile  $x$  can be decomposed in two steps: first allocate  $t'$  units, where  $t' \leq t$ , yielding the shares  $y'$ , next allocate the remaining  $t - t'$  units at the demand profile  $x - y'$ .

In the scheduling interpretation, LC says that the sequence  $r(x)$  is “history independent”: for any  $t$ , the subsequence starting at  $t$  only depends upon the number of jobs for each agent unserved at that time; it does not depend on the particular scheduling of jobs up to  $t - 1$ .

Upper Composition (UC)  $r(t, x) = r(t, r(x_N - 1, x))$ , where  $t < x_N$

The allocation of  $t$  units at profile  $x$  can be made in two steps: first allocate  $t'$  units, where  $t \leq t'$ , yielding the shares  $y'$ , next allocate  $t$  units at profile  $y'$ . In the scheduling format, UC states that the sequence  $r(x)$  is “future-blind”: for any  $t$  the subsequence up to time  $t$  only depends upon the total number of jobs awarded to each agent up to that time; it is independent of the subsequent scheduling.

The axioms Upper and Lower Composition are dual of each other. Consistency, on the other hand, is a self-dual property.

We conclude this section with two useful reformulations of LC and UC respectively. The obvious proof of Lemma 1 is omitted.

**Lemma 1** *Fix an arbitrary mapping  $\gamma$  from  $\mathbb{N}^N \setminus \{0\}$  into  $N$  and such that  $\gamma(x) = i \implies x_i > 0$ . There is a unique rationing method satisfying LC and such that*

$$r(1, x) = e_i \implies \gamma(x) = i \tag{2}$$

*There is a unique rationing method satisfying UC and such that*

$$r(x_N - 1, x) = x - e_i \implies \gamma(x) = i \tag{3}$$

*These two methods are dual of each other.*

Thus a method meeting LC (resp. UC) is entirely described by the way it allocates the first unit of resources (resp. the first unit of deficit).

## 4 Standard of gains, standard of losses

**Definition 1** *A (deterministic) standard of gains (in short s.g.) is a linear ordering<sup>11</sup>  $\succ$  of  $N \times (\mathbb{N} \setminus \{0\})$  such that  $(i, x_i) \prec (i, x_i + 1)$  for all  $i$ , all  $x_i = 1, 2, \dots$ . The corresponding rationing method satisfies Lower Composition and is defined (via (2)) as follows:*

<sup>11</sup>complete, transitive, antisymmetric

$$\gamma(x) = i \stackrel{\text{def}}{\succ} (i, x_i) \succ (j, x_j) \text{ for all } j \neq i \quad (4)$$

The standard  $\succ$  determines which agent has the strongest claim to the first unit of the commodity by ranking a demand  $x_i$  by agent  $i$  versus a demand  $x_j$  by agent  $j$ . After decreasing the demand of the winner of the first unit, the standard determines the strongest claim to the second unit and so on.

**Proposition 1** *A rationing method  $r$  is a standard of gains method (Definition 1) if and only if it meets Consistency, Lower Composition and Demand Monotonicity.*

All proofs are in the Appendix.

Standard of gains methods equitable in the sense of Equal Treatment Ex Post are easy to describe. The standard  $\succ$  must give precedence to a larger demand:  $(i, z+1) \succ (j, z)$  for all  $i, j$  and  $z$ . Indeed assume  $(j, z) \succ (i, z+1)$  and consider a demand  $x_i = x_j = z+1, x_k = 0$  otherwise. Then the first two units must go to agent  $j$ , a contradiction of ETEP. Conversely, any standard  $\succ$  for which a larger demand ensures a stronger claim, yields an equitable rationing method.

An equitable s.g. may order the pairs  $(i, z), i \in N$ , arbitrarily, and this order, denoted  $\sigma_z$ , may depend upon the level  $z = 1, 2, \dots$

Here is an example with  $N = \{1, 2, 3, 4\}$ : for  $z$  odd we posit  $(1, z) \succ (2, z) \succ (3, z) \succ (4, z)$  and for  $z$  even  $(2, z) \succ (4, z) \succ (1, z) \succ (3, z)$ . For instance at the profile of demands  $x = (3, 2, 2, 5)$ , the scheduling sequence is

$$r(x) = (4, 4, \underbrace{1, 4}_{\sigma_3}, \underbrace{2, 4, 1, 3}_{\sigma_2}, \underbrace{1, 2, 3, 4}_{\sigma_1},)$$

Reading this sequence backward (allocating successive units of deficits) amounts to enumerate the pairs  $(i, z)$  increasingly along the standard  $\succ$ , dropping an agent once he has appeared  $x_i$  times: here agents 2, 3 drop in the third round  $(i, 3)$  and agent 1 drops in the next one.

The example suggests an interesting interpretation of many s.g. methods equitable or not. Suppose that the standard  $\succ$  can be represented as an infinite sequence like:

$$\tilde{\alpha} = (i_1, z_1), (i_2, z_2), \dots, (i_k, z_k), \dots \quad k = 1, 2, \dots$$

where  $(i_{k+1}, z_{k+1}) \succ (i_k, z_k)$  and all pairs  $(i, z)$  in  $N \times \mathbb{N} \setminus \{0\}$  are listed. For instance  $(i_1, z_1)$  is the smallest pair of all and  $z_1 = 1$ . Note that not all standards can be so represented. An example is a fixed priority like  $1 > 2 > \dots > n$ , for which the standard is lexicographic;  $(i, x_i) \succ (j, x_j)$  iff  $i < j$  or  $i = j$  and  $x_i > x_j$ .

Because the standard  $\succ$  over the pairs  $(i, z)$  is increasing in  $z$ , the sequence  $\tilde{\alpha}$  is entirely determined by its projection  $\alpha$  on  $N$ , namely the *fixed path*:

$$\alpha = (i_1, i_2, \dots, i_k, \dots)$$

where each agent  $i$  appears infinitely many times.

To such a fixed path, we associate a scheduling method  $r$  as follows. For all  $x$ ,  $r(x)$  *read backwards* follows  $\alpha$ , skipping agent  $i$  once he has appeared  $x_i$  times. That is, the rationing method  $r$  distributes the successive units of deficit in the fixed order given by  $\alpha$ , skipping agent  $i$  after he appears  $x_i$  times. Clearly,  $r$  is the standard of gains method with the following standard  $\succ$ :

$$(i, x_i) \succ (j, x_j) \stackrel{def}{\iff} \{j \text{ appears } x_j \text{ times in } \alpha \text{ before } i \text{ appears } x_i \text{ times}\}$$

In the above example with 4 agents, our equitable standard of gains method is associated with the fixed path:

$$\alpha = (\underbrace{4, 3, 2, 1}, \underbrace{3, 1, 4, 2}, \underbrace{4, 3, 2, 1}, \underbrace{3, 1, 4, 2}, \dots)$$

An ordering  $\succ$  of  $N \times \mathbb{N} \setminus \{0\}$ , increasing in  $z$ , can be enumerated as an increasing sequence  $\tilde{\alpha}$  if and only if it is a well order, namely for every pair  $(i, z)$  the set of  $(j, z')$  below  $(i, z)$  is finite. Because  $\succ$  increases in  $z$ , this is equivalent to the following assumption:

$$\text{for all } i, j, i \neq j \text{ and all } x_i, \text{ there exists } x_j \text{ such that } (j, x_j) \succ (i, x_i) \quad (5)$$

The following statement summarizes the above discussion.

**Lemma 2** *Given an arbitrary fixed path  $\alpha$  (a sequence in  $N$  where each agent appears infinitely many times), the method  $r$  distributing successive units of deficit along  $\alpha$  is a standard of gains method satisfying the following archimedian property:*

*for all  $i, j, i \neq j$  and all  $x_i$ , there exists  $x_j$  such that  $r(1, x_i e_i + x_j e_j) = e_j$*

*Conversely, any standard of gains method with the archimedian property is a fixed path method.*

The archimedian property is a mild equity property, much weaker than ETEP<sup>12</sup>: any given demand of  $i$  can be “outweighed” by a large enough demand of  $j$ .

In the duality operation, standard of gains methods are turned into standard of losses methods. A (deterministic) standard of losses (s.l.) is the same mathematical object as a standard of gains, namely a linear ordering  $\succ$  of  $N \times (\mathbb{N} \setminus \{0\})$

<sup>12</sup>One checks easily that ETEP plus LC imply the archimedian property.

increasing in the second coordinate. The corresponding rationing method meets Upper Composition and is defined by (3) and the mapping  $\gamma$  defined in (4). Naturally, this method is the dual of the s.g. method of Definition 1. Therefore the next two results are mechanically adapted from Proposition 1 and Lemma 2.

**Proposition 1\*** *A rationing method is a standard of losses method if and only if it meets Consistency, Upper Composition and Demand Monotonicity\*.*

**Lemma 2\*** *Given an arbitrary fixed path  $\alpha$  (a sequence in  $N$  with infinitely many occurrences of each agent  $i$ ), the method  $r$  distributing successive units of the commodity along  $\alpha$  is a standard of losses method with the following archimedean property:*

*for all  $i, j, i \neq j$ , all  $x_i$ , there exists  $x_j$  such that  $r^*(1, x_i e_i + x_j e_j) = e_j$ .*

*Conversely, any standard of losses method with the above archimedean property is a fixed path method.<sup>13</sup>*

The s.l. methods equitable in the sense of Equal Treatment Ex Post distribute successive units along a fixed path  $\alpha$  consisting of successive orderings of  $N$  (i.e., each agent appears  $k$  times in the first  $nk$  terms of the sequence). Each such method is a deterministic version of fair queuing.

**Remark 1** As described above a fixed priority method relies on a given ordering  $\sigma$  of  $N$ , where  $\sigma(1) = i$  is the agent with the highest priority and so on. Then the scheduling sequence is as follows:

$$r(x) = \underbrace{(\sigma(1), \dots, \sigma(1))}_{x_1}, \underbrace{(\sigma(2), \dots, \sigma(2))}_{x_2}, \dots, \underbrace{\dots, \sigma(n)}_{x_n}$$

This method is both a s.g. and a s.l. method. There are no other methods at the intersection of these two families. This can be checked directly from the definitions. It also follows from Theorem 1 in Moulin [2000] characterizing the set of fixed priority methods by the combination of CSY, LC and UC. In particular, a s.g. (resp. s.l.) method satisfies UC (resp. LC) if and only if it is a fixed priority method.

**Remark 2** Proposition 1 (and Proposition 1\*) are tight results.

Drop Demand Monotonicity and the following “ultraprogressive” methods (among many others) emerge. Fix an ordering of  $N$ , say  $1 > 2 > \dots > n$ . For any profile of demands  $x$ , let  $\sigma_x$  be the ordering where a lower demand means a higher priority, ties being broken in favor of the smaller index agent: e.g.  $x = (3, 2, 5, 2)$  yields  $\sigma_x : 2 > 4 > 1 > 3$ . Our method allocates  $x$  according to the priority  $\sigma_x$ . It fails DM because raising my demand may lower my priority status. One checks easily CSY and LC.

<sup>13</sup>Fixed path methods emerge also in the (different) model of fair division with singlepeaked preferences. They are characterized there by incentive compatibility and consistency: see Barbera, Jackson and Neme [1997] and Moulin [1999b].

Any standard of gains (resp. losses) method meets DM as well as DM\*, although the latter property is not needed for the characterization result. Therefore if we drop the LC requirement, any s.l. method meets the other two requirements of Proposition 1.

Finally we drop CSY and we check that the pair of axioms LC + DM contains (much) more than the s.g. methods, except in the case  $|N| = 2$  (where CSY is vacuous). For any partition of  $N$  as  $N = \cup N_k$ , with  $k = 1, \dots, K$ , we can define the composition of  $K + 1$  methods  $r_0, r_1, \dots, r_K$ . Given a problem  $(t, x)$ , we use first  $r_0$  to divide the  $t$  units among the  $K$  subgroups  $N_k$ , taking  $x_{N_k}$  to be  $N_k$ 's demand; then we use  $r_k$  to allocate  $N_k$ 's share between the agents in  $N_k$ . The composition respects the properties LC and DM (as well as UC and DM\*) but not CSY. For instance if  $r$  is a two person equitable s.g. method (associated, say, with the fixed path  $1, 2, 1, 2, 1, \dots$ ) and if  $N = \{1, 2, 3, 4\}$  is partitioned as  $\{1, 2\}, \{3, 4\}$ , the composition of  $r, r, r$ , fails Consistency.

## 5 The probabilistic model

A probabilistic rationing method  $r$  associates to every deterministic rationing problem  $(t, x)$  a random vector of shares  $Y = r(t, x)$ , such that  $0 \leq Y_i \leq x_i$  for all  $i$  and  $Y_N = t$ . As this will cause no confusion, we often identify the random variable  $Y$  and its probability distribution on  $\mathbb{N}^N$ .

If a given vector  $a$  in  $\mathbb{N}^N$  is an upperbound on all demand profiles (so that  $0 \leq x_i \leq a_i$  for all  $i$ ), we speak of a rationing method  $r^a$  restricted to  $[0, a]$ . This notation is useful below in the description of probabilistic standard of gains methods. The duality operation is defined and interpreted exactly as in the deterministic model: property (1).

We postpone until after Lemma 3 below the discussion of the scheduling interpretation for probabilistic rationing methods. See in particular Remark 3.

In the probabilistic model, we have two different and logically unrelated equity axioms. Equal Treatment Ex Post is the same as in the deterministic model: if  $x_i = x_j$  then  $|y_i - y_j| \leq 1$  for every realization  $y$  of the random variable  $Y$ . The new axiom is:

Equal Treatment Ex Ante (ETEA):  $x_i = x_j \Rightarrow Y_i \sim Y_j$ .

(these two random variables have identical probability distributions.)

The proportional method below is an example of a method meeting ETEA but not ETEP. A deterministic method fails ETEA but can meet ETEP.

The five axioms driving the characterization results in the deterministic model extend straightforwardly to the probabilistic one. Their duality relations are preserved. Denote by  $\leq$  the stochastic dominance relation between two random variables  $Z, Z'$  on  $\mathbb{N}$ :

$$Z \leq Z' \stackrel{def}{\Leftrightarrow} \text{prob}\{Z \geq z\} \leq \text{prob}\{Z' \geq z\} \text{ for all } z \in \mathbb{N}$$

Demand Monotonicity (DM):  $Y_i = r_i(t, x) \leq Y'_i = r_i(t, x + e_i)$

Demand Monotonicity\*(DM\*):  $Y'_i = r_i(t + 1, x + e_i) \leq Y_i + 1 = r_i(t, x) + 1$

Consistency says that sharing  $t$  units given the demands  $x$  can be done in one shot or can be decomposed in two stages: first compute agent  $j$ 's share (a realization  $y_j$  of the random variable  $Y_j = r_j(t, x)$ ), next divide randomly  $t - y_j$  units among  $N \setminus j$  given the demands ( $x \upharpoonright^j 0$ ). The important point is that both random draws are stochastically independent:

Consistency (CSY)<sup>14</sup>:  $r_{-j}(t, x) = r_{-j}^2(t - r_j^1(t, x), (x \upharpoonright^j 0))$  for all  $j$ .

where  $r_{-j}(t, x)$  is the protection of  $r(t, x)$  on  $N \setminus j$  and the superscripts  $r^1$  and  $r^2$  remind us that these two random variables are stochastically independent. An alternative formulation of CSY uses the probability distribution of the random variable  $r(t, x)$ . If  $r(t, x)(y)$  is the probability that the realization of  $r(t, x)$  is  $y$ , the CSY axiom takes the form:

$$r(t, x)(y) = r(t, x)(Y_j = y_j) \cdot r(t - y_j, (x \upharpoonright^j 0))(y \upharpoonright^j 0) \text{ for all } t, x, y, j.$$

Lower Composition (LC):  $r(t, x) = r^1(1, x) + r^2(t - 1, x - r^1(1, x))$ .

Upper Composition (UC):  $r(t, x) = r^2(t, r^1(x_N - 1, x))$ .

Here again,  $r^1$  and  $r^2$  denote two stochastically independent random variables. In the more explicit formulation with probability distributions:

$$LC : r(t, x)(y) = \sum_i r(1, x)(e_i) \cdot r(t - 1, x - e_i)(y - e_i) \text{ for all } t \geq 1, x, y$$

(sum over those  $i$  s.t.  $x_i > 0$ )

$$UC : r(t, x)(y) = \sum_i r(x_N - 1, x)(x - e_i) \cdot r(t, x - e_i)(y) \text{ for all } t < x_N, x, y$$

(we set  $r(t, x - e_i)(y) = 0$  if  $y_i = x_i$ ).

As in the deterministic case, the LC (resp. UC) axiom means that a method is entirely determined by the way it allocates the first unit of the commodity (resp. the first unit of deficit).

**Lemma 3** *Fix an arbitrary mapping  $p$ , associating to every non zero profile of demands  $x$  a probability distribution  $p(x)$  over its support, and such that:  $p_i(x) > 0 \Rightarrow x_i > 0$ .*

*There is a unique probabilistic rationing method satisfying LC (resp. UC) and such that:*

---

<sup>14</sup>In the deterministic model – but not the probabilistic one – the equality of  $r_{-j}$  and  $r_{-j}^2$  is equivalent to that of  $r_i$  and  $r_i^2$ , for all  $i \neq j$ . Hence the definition of CSY in Section 3 is the same.

$$r(1, x)(e_i) = p_i(x) \quad (\text{resp. } r(x_N - 1, x)(x - e_i) = p_i(x))$$

These two methods are dual of each other.

The LC method associated with  $p$  works by allocating the first unit according to the probability distribution  $p(x)$ , the second unit according to  $p(x - e_i)$  where  $i$  is the recipient of the first unit, and so on. Therefore given the demand profile  $x$ , the sequence of random variables  $r(1, x), \dots, r(t, x), \dots, r(x_N - 1, x)$  is equivalently described as a *probabilistic scheduling* sequence  $r(x) = (i_1, \dots, i_t, \dots, i_{x_N})$ , where given  $i_1, \dots, i_{t-1}$ , the probability distribution of  $i_t$  is  $p(x - (e_{i_1} + \dots + e_{i_{t-1}}))$ . That is to say for any given  $t$  the probability distribution of  $e_{i_1} + \dots + e_{i_t}$  is precisely  $r(t, x)$ .

Thus any LC rationing method can be described as a random scheduling process, and by duality the same is true for any UC rationing method<sup>15</sup>. In this representation, Lower Composition means that the probability distribution of  $i_t$  only depends upon the remaining (unfulfilled) demands after the realization of  $i_1, \dots, i_{t-1}$ : this is “history independence” as in Section 3. Upper Composition is similarly interpreted as “future-blindness”. Consistency means that the (probability distribution of the) random  $(i, j)$  - subsequence of  $r(x)$  only depends upon the demands  $x_i, x_j$ , not upon the demands of other agents.

We give now more details on the non trivial relationship between rationing methods and probabilistic scheduling methods.

**Remark 3** In the deterministic model, a resource monotonic rationing method ( $t \leq t' \Rightarrow r(t, x) \leq r(t', x)$ ) can be interpreted as a scheduling method, because the increments of the sequence  $r(1, x), r(2, x), \dots$  are the coordinate vectors  $e_{i_1}, e_{i_2}, \dots$ . In the probabilistic model, Resource Monotonicity means that  $r_i(t, x)$  is stochastically dominated by  $r_i(t + 1, x)$  for all  $i, t$  and  $x$ . A probabilistic scheduling method associates to any  $x$  a random sequence  $\{i_1, \dots, i_{x_N}\}$  where each agent  $i$  appears exactly  $x_i$  times. The corresponding random variables  $e_{i_1} + \dots + e_{i_t}$  define a resource monotonic probabilistic rationing method. But the converse statement is not true, except in the two agent case ( $|N| = 2$ ). An example below establishes this claim.

Fix a rationing method  $r$ , a rationing problem  $(t, x)$ ,  $t > 0$ , and a set  $A$  of feasible allocations:

$$A \subseteq \{y \mid 0 \leq y_i \leq x_i \text{ for all } i \text{ and } y_N = t\}$$

Denote by  $A^{-1}$  the set of its predecessors in the problem  $(t - 1, x)$ :

$$A^{-1} = \{y \mid 0 \leq y_i \leq x_i \text{ for all } i \text{ and for some } j : y + e_j \in A\}$$

---

<sup>15</sup>Given  $i_{x_N}, \dots, i_{t+1}$ , the probability distribution of  $i_t$  is  $p(x - (e_{i_{x_N}} + \dots + e_{i_{t+1}}))$ .



If  $r$  can be deduced from a probabilistic scheduling method, any sequence  $\{i_1, \dots, i_{x_N}\}$  passing through  $A$  must pass through  $A^{-1}$  as well, therefore the inequality:

$$r(t, x)(A) \leq r(t-1, x)(A^{-1}) \quad (6)$$

is a necessary condition for such a representation of  $r$ . Conversely, if this inequality holds for all  $t, x$  and  $A$ , it is easy to show that  $r$  is the “projection” of a probabilistic scheduling method.<sup>16</sup>

We give now an example of a resource monotonic rationing method violating inequality (6). We set  $x = (3, 3, 3)$  and define  $r(t, x)$  for  $t = 5, 6$ :

$$\begin{aligned} r(5, x) &: (0, 2, 3), (3, 0, 2), (2, 3, 0) \text{ each with proba. } 1/9 \\ &\quad (2, 2, 1), (2, 1, 2), (1, 2, 2) \text{ each with proba. } 2/9 \\ r(6, x) &: (2, 1, 3), (1, 3, 2), (3, 2, 1) \text{ each with proba. } 1/9 \\ &\quad (2, 2, 2) \text{ with proba. } 2/3 \end{aligned}$$

Resource Monotonicity follows because

$$\begin{aligned} r_i(5, x) &= 1, 2 \text{ or } 3 \text{ with respective probas } 2/9, 5/9, 1/9 \\ r_i(6, x) &= 1, 2 \text{ or } 3 \text{ with respective probas } 1/9, 7/9, 1/9 \end{aligned}$$

One can even check a stronger property than RM namely for all  $z$  :

$$r(5, x)(Y \geq z) \leq r(6, x)(Y \geq z) \text{ and } r(5, x)(Y \leq z) \geq r(6, x)(Y \leq z)$$

However, let  $A$  be the support of  $r(6, x)$  namely  $A = \{(2, 2, 2), (2, 1, 3), (1, 3, 2), (3, 2, 1)\}$  so that  $A^{-1}$  does not contain  $(0, 2, 3), (3, 0, 2)$  or  $(2, 3, 0)$ : thus inequality (6) fails.

## 6 Three basic methods and their characterization

We define successively the proportional, fair queuing and fair queuing\* methods. They are the counterpart in the probabilistic model of the proportional, uniform gains and uniform losses methods of the classical model with divisible goods and deterministic allocations. All three methods are equitable in the ex ante sense.

### *Proportional*

---

<sup>16</sup>This follows easily from Hall’s theorem, see, e.g., Bazaraa et al.[1990]. We thank Rakesh Vohra for pointing out this result.

The proportional method follows in Lemma 3 from the proportional probability distribution  $p_i(x) = x_i/x_N$ . It is the corresponding LC method, and the UC method as well; in particular it is a self-dual method. Its interpretation as a probabilistic scheduling method is especially simple: given the demand profile  $x$ , we place  $x_i$  balls labeled  $i$  in an urn, for  $i = 1, 2, \dots, n$ . The scheduling sequence obtains by drawing successively all balls from the urn without replacement (equivalently, all sequences where each  $i$  appears  $x_i$  times are equally probable). In particular, the expected shares are “classically” proportional to individual demands. The easy proof of the above facts is omitted, see Moulin [1999a].

The proportional method meets all six axioms introduced in the previous section. As already noticed in Section 2, it does not meet Equal Treatment Ex Post.

#### *Fair Queuing*

This method allocates one unit to each agent in successive round robin fashion, dropping an agent only when his demand is met. To be precise, the method first selects at random and with uniform probability an ordering of the  $m_1$  agents with a positive demand, and allocates the first  $m_1$  units in that order; it draws next an ordering of the  $m_2$  agents whose remaining demand is positive and so allocates the next  $m_2$  units; and so on.

This definition makes clear that fair queuing violates Lower Composition: the recipient  $i_2$  of the second unit is selected with uniform probability among  $N_1 \setminus \{i_1\}$ , where  $N_1$  is the set of agents with a positive demand and  $i_1$  received the first unit: knowledge of the vector  $x - e_1$  is not enough to determine the scheduling process from time 2 onward.

An equivalent definition of fair queuing describes how it allocates successive units of deficit. This formulation makes clear that it meets Upper Composition. Fair queuing distributes the first unit of deficit with uniform probability among all agents with the largest demand: see Figure 1.

$$p_i(x) = \frac{1}{n(x)} \text{ if } i \in N(x); = 0 \text{ if } i \notin N(x)$$

$$\text{where } i \in N(x) \text{ iff } \{x_j \geq x_i \text{ for all } i\}, n(x) = |N(x)| \quad (7)$$

The third definition of fair queuing makes the probability distribution  $r(t, x)$  explicit. It relies on the uniform gains method of the classical model. To a rationing problem  $(t, x)$  we associate the real number  $\lambda$  defined by  $\sum_i \min\{\lambda, x_i\} = t$ . Let  $k$  be the largest integer bounded above by  $\lambda$  and  $M(t, x)$  be the set of agents  $i$  such that  $\lambda < x_i$  ( $M(t, x)$  is non empty when  $t < x_N$ ), with cardinality  $m(t, x)$ . Notice that  $m(t, x) \cdot (\lambda - k) = q$  is an integer (possibly zero). Now the distribution of  $r(t, x)$  is as follows:

give  $\min\{k, x_i\}$  units for sure to  $i$ , for all  $i$ , and spread  $q$  units over  $M(t, x)$  with uniform probability (giving at most one unit per agent)

See Figure 1, in the case  $n = 2$ . This figure shows that fair queuing is closely related to the family of equitable (in the ex post sense) deterministic standard of losses methods: drawing at random and uniformly a scheduling sequence from that family is a representation of fair queuing as a probabilistic scheduling method. This observation holds true for any number of agents.

The equivalence of the three above definitions of fair queuing is proven in Moulin [1999a]; it is easy to check directly. In particular, the expected shares are equal to the uniform gains shares of the classical model, namely  $E[r(t, x)] = \min\{\lambda, x_i\}$ .

The fair queuing method meets all axioms discussed so far, except Lower Composition. In particular, it is equitable both in the ex ante (ETEA) and ex post (ETEP) sense.

#### *Fair Queuing\**

The fair queuing\* method is the dual of fair queuing. Therefore it is the LC method associated via Lemma 3 with the probability distribution (7). Alternatively this method allocates successive units of deficit by means of the round robin distributions discussed above. See Figure 2 in the case of  $n = 2$ , showing the probabilistic scheduling interpretation of fair queuing\*: it picks at random and uniformly among all equitable deterministic standard of gains methods. Fair queuing\* meets all the axioms introduced so far except Upper Composition.

The Consistency Property is crucial to the characterization of the rich family of standard of gains/losses methods in the deterministic (Propositions 1, 1\* above) and probabilistic (Theorem 2 below) models. Remarkably, it plays no role in that of our three basic methods.

**Theorem 1** *a) The proportional method is the only probabilistic rationing method meeting Lower and Upper Composition, and Equal Treatment Ex Ante.*

*b) Fair Queuing is the only method meeting Upper Composition, Demand Monotonicity\*, Equal Treatment Ex Ante and Equal Treatment Ex Post.*

*c) Fair Queuing\* is the only method meeting Lower Composition, Demand Monotonicity, Equal Treatment Ex Ante and Equal Treatment Ex Post.*

Statement a is Theorem 1 in Moulin [1999a]. The proof of statements b and c is in the Appendix. All three statements are tight characterization results: see the discussion in Section 8.

## 7 Probabilistic standard of gains/losses

We describe a rich family of methods containing equitable as well as inequitable methods (for instance the fixed priority methods). We start by the somewhat special case of a two agent problem.

**Lemma 4** *Let  $z = z_i e_i + z_j e_j$ ,  $i \neq j$ ,  $z_i, z_j > 0$ , be an element of  $\mathbb{N}^N$  of dimension 2 and  $r^z$  be a rationing method restricted to  $[0, z]$ . Then  $r^z$  meets*

LC and DM if and only if it is associated (via Lemma 3) with a (restricted) probability distribution  $p^z$  such that:

$p_i^z(x_i e_i + x_j e_j)$  is non decreasing in  $x_i$  and non increasing in  $x_j$ , for  $0 \leq x \leq z$ .

**Definition 2** A (probabilistic) standard of gains is a preordering<sup>17</sup> of  $\mathbb{N} \times (\mathbb{N} \setminus \{0\})$ , non decreasing in its second coordinate:

$$(i, x_i) \succsim (i, x_i + 1) \text{ for all } i, \text{ all } x_i = 1, 2, \dots$$

An indifference class of the standard of gains  $\succsim$ , denoted  $\delta$ , takes the form

$$\delta = \bigcup_{i \in S} \{i\} \times I_i \text{ where } I_i = [a_i, b_i] \text{ or } [a_i, \infty[ \text{ is an interval of } \mathbb{N} \setminus \{0\}$$

The set of agents  $S$  is called the *support* of  $\delta$  and its cardinality is the *type* of  $\delta$ . The *depth* of  $\delta$  is the vector  $\theta(\delta) = z$  where  $z_i = |I_i|$  (the cardinality of  $I_i$ ) for all  $i$ .

Given a profile of demands  $x, x \in \mathbb{N}^N \setminus \{0\}$ , we let  $\delta(x)$  be the indifference class intersecting  $\bigcup_i (i, x_i)$  and highest in  $\succsim$ , and define the *depth* of  $x$  in  $\delta(x)$  as the following vector  $d(x)$ :

$$d(x)_i = 0 \text{ if } (i, x_i) \notin \delta(x), = x_i - a_i + 1 \text{ if } (i, x_i) \in \delta(x)$$

where  $a_i$  is the smallest integer s.t.  $(i, a_i) \sim (i, x_i)$

Notice that if  $(i, x_i)$  is the unique maximum of  $\succsim$  among  $(j, x_j)$ , the depth  $d(x)$  is of dimension 1 and borne by  $x_i$ . If  $\delta(x)$  is of type 1, the maximum of  $\succsim$  among  $(j, x_j)$  is unique but the converse is not true: for instance if  $x = 3e_i$ , the unique maximum of  $\succsim$  is  $(i, 3)$  but the support of the indifference class of  $(i, 3)$  may contain other agents as well.

**Definition 3** Standard of gains methods

Given are a standard of gains  $\succsim$  and, for each indifference class  $\delta$  of type 2, a rationing method  $r^z$  on  $[0, z]$ , where  $\theta(\delta) = z$ , satisfying LC and DM, and with probability distribution  $p^z$  (Lemma 4).

The corresponding probabilistic standard of gains method is the LC method associated (via Lemma 3) with the following probability distribution  $p$ . For all non null demand profile  $x$ :

$$\begin{aligned} p_i(x) &= 1 \text{ if } (i, x_i) \succ (j, x_j) \text{ for all } j \neq i \\ p(x) &= p^z(y) \text{ if } \delta(x) \text{ is of type 2, } z = \theta(\delta(x)) \text{ and } y = d(x) \\ p(x) &= \frac{1}{y_N} y \text{ if } \delta(x) \text{ is of type 3 or more and } y = d(x) \end{aligned}$$

---

<sup>17</sup>complete, transitive

The definition makes sense because if  $\delta(x)$  is of type 2 or more and  $(i, x_i)$  is the unique maximum of  $\succ$ , the depth  $y = d(x)$  is borne by  $e_i$  hence  $p_i(y) = 1$  under any rationing method.

Given a demand profile  $x$ , the above method identifies the depth of  $x$  in its indifference class,  $y = d(x)$ , and allocates the first  $y_N$  units of gains as follows:

- all units go to agent  $i$  if  $(i, x_i)$  is the unique maximum of  $\succsim$
- use  $r^z$  at the profile  $y$  if  $\delta(x)$  is of type 2 and  $z = \theta(\delta(x))$
- use the proportional method at profile  $y$  if  $\delta(x)$  is of type 3 or more.

After  $y_N$  units have been so allocated, the remaining demand profile is  $x' = x - y$ . By construction,  $\delta(x')$  is an indifference class below  $\delta(x)$  and we repeat the process to allocate the next  $y'_N$  units,  $y' = d(x')$ .

**Remark 4** *There is some flexibility in the representation of a probabilistic rationing method by a standard of gains in Definition 3. We can sometime alter the pattern of indifference classes of type 1 and 2 without changing the rationing method. Two adjacent indifference classes of type 1 involving the same agent can be concatenated into one indifference class; two adjacent indifference classes with total support involving only two agents ( $S_1 \cup S_2 = \{i, j\}$ ) can similarly be merged into a single class of type 2 (provided we concatenate the corresponding methods  $r^z$ ).*

In the rest of this section, we illustrate the rich family of s.g. methods, as well as its dual, the standard of losses methods.

The proportional method corresponds to the preordering  $\succsim$  for which all pairs  $(i, x_i)$  are indifferent. The fair queuing\* method corresponds to the finest symmetric preordering  $\succsim$ , namely:

$$(i, x_i) \succsim (j, x_j) \stackrel{def}{\iff} x_i \geq x_j$$

Here the indifference class  $\delta(x)$  contains the pairs  $(i, \max_j x_j)$  for all  $i$ , and the depth  $d(x)$  is the sum of the unit vectors  $e_i$  over all agents with largest demand. Fair queuing\* gives equal probabilities of receiving first unit to all agents with the largest demand: this is, as in Definition 3, the proportional method applied at  $d(x)$ .

The interesting subset of standard of gains methods equitable in the ex ante sense is now described. The preordering  $\succsim$  is *symmetric* if  $x_i = x_j \Rightarrow (i, x_i) \sim (j, x_j)$ . All indifference classes are then of type  $n$ , and the preordering is entirely described by a finite or infinite partition of  $\mathbb{N} \setminus \{0\}$  in “indifference intervals”  $I_q = [b_{q-1}, b_q]$ , where  $b_0 = 0$  and  $b_q < b_{q+1}$ . Of course we set  $b_Q = +\infty$  if the partition is finite and  $I_Q$  is its highest interval.

Given the sequence  $b_q$ , we write for all positive integer  $\gamma(a) = q$  iff  $b_q < a \leq b_{q+1}$ , so that  $(i, x_i) \succsim (j, x_j)$  iff  $\gamma(x_i) \geq \gamma(x_j)$ . In our next statement we use the following notations:  $e = (1, 1, \dots, 1)$  and  $(a \wedge b)_i = \min\{a_i, b_i\}$ .

**Lemma 5** *i) Assume  $N$  contains three or more agents. To any (finite or infinite) increasing sequence  $b_q$  such that  $b_0 = 0$  and  $\sup_q b_q = +\infty$ , we associate a standard of gains method meeting Equal Treatment Ex Ante as follows. Fix a demand profile  $x$  such that  $\max_i \gamma(x_i) = q^*$ , and write  $x^q = x \wedge (b_q e)$  for  $q = 1, \dots, q^*$  :*

- *allocate proportionally the first  $x_N - x_N^{q^*}$  units at the profile  $x - x^{q^*}$*
- *allocate proportionally the next  $x_N^{q^*} - x_N^{q^*-1}$  units at the profile  $x^{q^*} - x^{q^*-1}$*
- *...*
- *allocate proportionally the last  $x_N^1$  units at the profile  $x^1$*

*Conversely, a standard of gains method meeting ETEA is constructed as above.*

*ii) If  $N = \{1, 2\}$  a similar statement holds, except that in each block  $]b_{q-1}e, b_qe]$  we can use an arbitrary symmetric probability distribution  $p^z$  as in Lemma 4.*

To illustrate the above result, we give an example with four agents,  $N = \{1, 2, 3, 4\}$  and demands  $x_i$  are distributed as follows within the indifference classes

$$\gamma(x_1) = 3, \quad \gamma(x_2) = 5, \quad \gamma(x_3) = \gamma(x_4) = 1$$

- the first  $(x_2 - b_4)$  units are given to agent 2
- the next  $(x_1 + b_4 - 2b_3)$  units are allocated proportionally between 1 and 2 with initial demands  $y_1 = x_1 - b_3, y_2 = b_4 - b_3$
- the next  $2(b_3 - b_2)$  units are allocated proportionally between 1 and 2 with initial demands,  $y_i = b_3 - b_2, i = 1, 2$
- the next  $(x_3 + x_4 + 2b_2 - 4b_1)$  units are allocated proportionally with initial demands  $y_i = b_2 - b_1$  for  $i = 1, 2, y_j = x_j - b_1$  for  $j = 3, 4$
- the last  $4b_1$  units are allocated proportionally with initial demands  $y_i = b_1$  for all  $i$

Let us read the random scheduling sequence backward in the above example, that is to say we look at the allocation of successive units of deficit. Because the proportional method is self-dual, the above algorithm is simply reversed: the first  $4b_1$  units are allocated proportionally, with initial demand  $y_i = b_1$  for all  $i$ ; and so on.

This is the probabilistic counterpart of the representation of a deterministic standard of gains method by a fixed path along which successive units of deficit are handed out: Lemma 2. The fixed path generalizes to a “block-path”, with

successive blocks  $]b_q e, b_{q+1} e]$ . The random scheduling sequence follows a proportional sequence in each block moving upward from  $]0, b_1 e]$ , and restricting individual demands in each block to the length of the block or the remaining demand, whichever is smallest. Figure 3 shows an example with two agents: notice the similarity with Figure 2 for Fair Queuing\*.

The above description applies to all standard of gains methods equitable ex ante. For non equitable standard of gains methods, a similar representation along a fixed (asymmetric) block-path can be given, provided a certain archimedian property is satisfied: this probabilistic counterpart of Lemma 2 is the subject of Remark 5 below.

Like in the deterministic case, the duality operation transforms standard of gains into standard of losses methods. Fix a preordering of  $N \times (\mathbb{N} \setminus \{0\})$  as in Definition 2, now called a standard of losses, and use it as in Definition 3 to construct a probability distribution  $p(x)$  on  $N$  for all  $x$ . The corresponding UC method (Lemma 3) is the standard of losses method associated with our preordering. Of course for each indifference class of type 2 we must provide a rationing method meeting UC and DM\*, and described by a probability distribution  $p^z$  as in (the dual of) Lemma 4.

The prominent examples of such methods are the proportional and fair queuing methods. More generally all standard of losses methods equitable ex ante are described by the dual of Lemma 5. Their random scheduling sequences can be described backward by the algorithm of Lemma 5 (now interpreted as distributing successive units of deficit) or forward by means of the fixed block-path described three paragraphs above. The latter description is more natural in the scheduling context: for the sake of comparison with Lemma 5, we describe the corresponding algorithm.

Fix an increasing sequence  $b_q$  as in Lemma 5 and assume  $|N| \geq 3$  for brevity. Given a demand profile  $x$  we write  $x^q = x \wedge (b_q e)$ , and set  $\max_i \gamma(x_i) = q^*$  so that  $x^{q^*+1} = x$  :

- allocate proportionally the first  $x_N^1$  units at profile  $x^1$ .
- allocate proportionally the next  $(x_N^2 - x_N^1)$  units at profile  $(x^2 - x^1)$
- ...
- allocate proportionally the last  $(x_N - x_N^{q^*})$  units at profile  $(x - x^{q^*})$

Comparing the two algorithms, above and in Lemma 5, shows why a standard of losses typically fails LC. After the first unit is allocated, say to agent  $i$ , his reduced demand  $x_i^1 - 1$  (in the first block) results in a lower chance of the next unit. But if  $x_i \geq b_1 + 1$ , the remaining overall demand  $x' = x - e_i$  has  $x' \wedge (b_1 e) = x \wedge (b_1 e) = x^1$  where the chance for  $i$  to get the first unit did not drop.

Figure 4 illustrates an equitable s.l. method with two agents. It is similar to Figure 1 depicting fair queuing.

**Remark 5** We discuss the possibility to represent a s.g. method (resp. s.l. method) by means of a fixed block-path, namely the probabilistic counterpart of Lemma 2 (resp. 2\*). Consider the following archimedean property of a (probabilistic) rationing method:

for all  $i, j, i \neq j$ , all  $x_i$ , there is  $x_j$  s.t.  $r(1, x_i e_i + x_j e_j)(e_j) > 0$   
 (resp.  $r^*(1, x_i e_i + x_j e_j)(e_j) > 0$ )

If  $r$  is a s.g. (resp. a s.l.) method, the archimedean property amounts to the same property (5) as in the deterministic case. This holds if and only if the indifference classes of  $\succsim$  can be increasingly enumerated as a (finite or infinite) sequence  $\delta^1, \dots, \delta^q, \dots$ . We write the indifference class  $\delta^q$  as:

$$\delta^q = \bigcup_N \{i\} \times ]b_i^{q-1}, b_i^q] \text{ where } b_i^{q-1} = b_i^q \text{ if } i \text{ is not in the support of } \delta^q$$

In the above definition, we set  $b^0 = 0$ . By construction we have  $\sup_q b_i^q = +\infty$ , whether the number of indifference classes is finite or infinite. The sequence  $\delta^1, \dots, \delta^q, \dots$  is thus described by the increasing sequence  $b^1, \dots, b^q$ .

We now generalize Lemma 5 as follows. Fix a sequence in  $\mathbb{N}^N$   $b^0, \dots, b^q, \dots$  such that  $b^0 = 0, b^q \leq b^{q+1}$ , and  $\sup_q b_i^q = +\infty$ , and for each block of dimension

2, fix a rationing method  $r^q$  restricted to  $[0, b^q - b^{q-1}]$ , as in Lemma 4. We write  $\gamma_i(x_i) = q$  iff  $b_i^q < x_i \leq b_i^{q+1}$ .

We define an archimedean standard of gains method as follows. Fix a profile  $x$  and write  $\max_i \gamma_i(x_i) = q^*$ ,  $x^q = x \wedge b^q$ :

- allocate the first  $x_N - x_N^{q^*}$  units at profile  $x - x^{q^*}$ : proportionally if  $b^{q^*+1} - b^{q^*}$  is not of dimension 2; according to  $r^{q^*+1}$  otherwise.
- allocate the next  $x_N^{q^*} - x_N^{q^*-1}$  units at profile  $x^{q^*} - x^{q^*-1}$ : proportionally if  $b^{q^*} - b^{q^*-1}$  is not of dimension 2; according to  $r^{q^*}$  otherwise.
- and so on.

Conversely, every archimedean s.g. method is constructed as above from a sequence  $b^q$ , together with a method  $r^q$  every time that  $b^q - b^{q-1}$  is of dimension 2. We omit for brevity the dual statement about archimedean standard of losses methods.

## 8 Main characterization results

**Theorem 2** *i) Every standard of gains method (Definition 3) meets the three properties Consistency, Lower Composition and Demand Monotonicity. Conversely a probabilistic rationing method meeting these three properties is a standard of gains method.*

*ii) The set of standard of losses method is similarly characterized by the combination of Consistency, Upper Composition and Demand Monotonicity\*.*

**Corollary 3** *The set of standard of gains methods described in Lemma 5 is characterized by the combination of the four axioms: Consistency, Lower Com-*



position, Demand Monotonicity and Equal Treatment Ex Ante. (We omit the straightforward dual statement).

Both statements in Theorem 2 are tight results for the same reason that Propositions 1 and 1\* are tight: Remark 2 in Section 4.

We check that the Corollary is tight. Drop Consistency and the *equal chance of gains* method<sup>18</sup> meets all three other axioms: this is the LC method such that  $p(x)$  is uniformly distributed among all agents with positive demand. Drop Demand Monotonicity, and the following ultraprogressive method emerges: it is the LC method where  $p(x)$  is uniformly distributed among all agents with smallest demand. Just like its deterministic counterparts (see Remark 2) this method meets CSY and LC, but fails DM. Tightness of the Corollary with respect to the LC and ETEA axioms is obvious.

Finally, we check that statement c (and, by duality, statement b) in Theorem 1 is tight.

Dropping LC allows Fair Queuing, dropping ETEP allows the proportional method, and dropping ETEA allows the deterministic s.g. methods equitable ex post (discussed just before Remark 1 in Section 4). Dropping DM allows the following variant of the *ultraprogressive* method above. Given a profile  $x$ , we let  $N^1(x)$  be the set of agents with the lowest demand. These agents are served first by Fair Queuing. Then the agents in  $N^2(x)$ , with the next lowest demand, are served in full by Fair Queuing. And so on.

## 9 Lower cum Upper composition

The set of rationing/scheduling methods meeting Lower and Upper Composition cannot be easily described. This is true even if we add the requirements of Demand Monotonicity and Demand Monotonicity\*, and is explained at the end of this section.

However, when we restrict attention to methods that are either deterministic, or fully random (in a sense made precise below) the impact of the combination LC+UC can be described.

a) *deterministic methods*

For any ordering  $\sigma$  of  $N$  (where  $\sigma(i) = 1$  reads  $i$  is first), we denote by  $prio(\sigma)$  the associated fixed priority method (Remark 1); we also write  $top(\sigma) = \sigma^{-1}(1)$  and  $bot(\sigma) = \sigma^{-1}(|N|)$ .

**Definition 4** For each nonempty subset  $S$  of  $N$ , choose an ordering  $\sigma_S$  of  $S$ , such that:

$top(\sigma_S) = i \Rightarrow \sigma_{S \setminus i}$  and  $\sigma_S$  coincide on  $S \setminus i$

$bot(\sigma_S) = j \Rightarrow \sigma_{S \setminus j}$  and  $\sigma_S$  coincide on  $S \setminus j$

The corresponding quasipriority method is defined as:

$r(x) = prio(\sigma_S)$  where  $S = \{i \in N \mid x_i > 0\}$  is the support of  $x$ .

---

<sup>18</sup>Moulin [1999a]

With two agents,  $|N| = 2$ , the two priority methods are the only quasipriority methods. With three agents, there are twelve quasipriority methods, six are the fixed priority methods and six are as follows:

$$\sigma_{123} : 1 > 2 > 3; \sigma_{13} : 3 > 1$$

Here agent 3 has priority over 1 if agent 2 is out, but otherwise agent 3 comes last after 1 and 2.

The number of quasipriority methods grows rapidly with  $|N|$ : there are 624 for four agents. The following facts (of which the proof is omitted for brevity) show that the set of quasipriority methods becomes complex as  $n$  grows. For any tournament on  $N$  (complete, antisymmetric relation) there is a quasipriority method that agrees with the tournament on pairs. On the other hand, there are quasipriority methods in which for any pair  $i, j$ , there is a subset  $S$  with  $i$  preceding  $j$  in  $\sigma_S$  and a subset  $S'$  with  $j$  preceding  $i$  in  $\sigma_{S'}$ <sup>19</sup>.

**Proposition 2** *A quasipriority method meets Lower and Upper Composition, Demand Monotonicity and Demand Monotonicity\*.*

*Conversely, a deterministic rationing/scheduling method meeting LC, UC, DM and DM\* is a quasipriority method.*

Proposition 2 is a negative result, because a quasipriority method is essentially as inequitable as a priority method: for instance, it violates the mild archimedean properties of Lemma 2 and 2\*, and, a fortiori, Equal Treatment Ex Post.

When we drop the DM and DM\* requirements, many more methods emerge, that combine strong inequity with unpalatable lack of monotonicity. For instance, assume  $N = \{1, 2, 3\}$  and fix a standard of gains  $\succ$  (Definition 1). Denote by  $prio(i, j, k)$  the priority method where  $i$  is served first, followed by  $j$  then  $k$ . We write  $i + 1$  (*resp.*  $i - 1$ ) with the convention  $3 + 1 = 1$  (*resp.*  $1 - 1 = 3$ ). The following method meets UC and LC:

$$\begin{aligned} \text{if } x_i > 0 \text{ for } i = 1, 2, 3 \text{ and } (i, x_i) \succ (j, x_j), (k, x_k), & \text{ use } prio(i - 1, i, i + 1) \\ \text{if } x_i = 0, & \text{ use } prio(i + 1, i - 1) \end{aligned}$$

Proposition 2 is a tight result. Drop DM\* and the following method in  $N = \{1, 2, 3\}$  meets LC, UC and DM:

$$\begin{aligned} \text{if } x_2 > 0 \text{ and } x_3 = 0, 1, \dots, 9 : & \text{ use } prio(1, 2, 3) \\ \text{if } x_2 = 0 \text{ and/or } x_3 = 10, 11, \dots, : & \text{ use } prio(2, 3, 1) \end{aligned}$$

---

<sup>19</sup>The first example occurs with 5 people ordered as:  $1 > 2 > 3 > 4 > 5$ ,  $5 > 1 > 2 > 3$ ,  $4 > 5 > 1 > 2$ ,  $3 > 4 > 5 > 1$ ,  $2 > 4 > 1$ ,  $4 > 1 > 3$ ,  $1 > 3 > 5$ ,  $3 > 5 > 2$ ,  $5 > 2 > 4$ , and all other orderings deduced from Definition 4.

b) *probabilistic methods*

We already know (statement a, Theorem 1) that the only LC+UC method equitable ex ante is the proportional method. Moreover, there is no LC+UC method equitable ex post<sup>20</sup>. We describe below a rich family of LC+UC methods that contains the proportional one and, as a limit case, the fixed priority methods.

We introduce two new axioms, conveying yet another principle of equity in our model: we use the notation  $E(Y_i)$  for the mean of the random variable  $Y_i$ :

Positive Shares (PS):  $\{x_i \geq 1, t \geq 1\} \Rightarrow \{E(r_i(t, x)) > 0\}$

Positive Risks (PR):  $\{x_i \geq 1, t \leq x_N - 1\} \Rightarrow \{E(r_i(t, x)) < x_i\}$

The two axioms are dual of each other. Positive shares says that an agent with a positive claim cannot be entirely left out of the distribution: he receives at least one unit with positive probability. Thus PS is an individual rationality condition. Dually, Positive Risk says that every agent with a positive demand must bear a share of the overall deficit  $x_N - t$ .

The proportional method meets PR and PS. Fair queuing is a PS method but it fails PR.

We define the family of *weighted proportional methods*, axiomatically characterized in Theorem 3 below. Every such method is described by a system of weights  $(w_{ij})$ , one for each pair  $(i, j), i \neq j$ :

$$\text{for all } i, j : w_{ij} > 0 \text{ and } w_{ij} + w_{ji} = 1$$

We need some notations to define our methods. Given a demand profile  $x$ , recall that a scheduling sequence  $s = \{i_1, i_2, \dots, i_{x_N}\}$  is such that each agent  $i$  appears exactly  $x_i$  times. We denote by  $\Delta(x)$  the set of such sequences. We write  $a_{ji}(s)$  for the number of times  $j$  precedes  $i$  in  $s$ , namely the number of pairs  $(t, t')$ ,  $1 \leq t < t' \leq x_N$ , such that  $i_t = j$  and  $i_{t'} = i$ . Note that  $a_{ij}(s) + a_{ji}(s) = x_i \cdot x_j$ . Finally we denote by  $P$  the set of ordered pairs  $(i, j)$  with  $i \neq j$  and by  $P^*$  the set of unordered pairs. Thus  $|P| = n(n-1)$  whereas  $|P^*| = n(n-1)/2$ .

Given a set of weights  $(w_{ij})$ , and a sequence  $s$  we define

$$w(s) = \prod_{(i,j) \in P} w_{ij}^{a_{ij}(s)} \text{ and } D(x) = \sum_{s \in \Delta(x)} w(s) \quad (8)$$

If  $x = x_i e_i$ , the set  $\Delta(x)$  contains a single sequence and  $w(s) = 1$ , so that  $D(x) = 1$  in this case. The following identity, where  $S(x)$  denotes the support of  $x$  ( $i \in S$  iff  $x_i > 0$ ), follows easily from (8):

$$D(x) = \sum_{j \in S(x)} \left( \prod_{i \in N \setminus j} w_{ij}^{x_i} \right) \cdot D(x - e_j) \quad (9)$$

---

<sup>20</sup>this is easy to check by looking at the case  $N = \{1, 2\}$  and the demand  $x = (2, 2)$ .

Therefore the following vector  $p(x)$  is a probability distribution over the support of  $x$ :

$$p_j(x) = \left( \prod_{i \in N \setminus j} w_{ij}^{x_i} \right) \cdot \frac{D(x - e_j)}{D(x)} \quad \text{for all } j \in S(x) \quad (10)$$

We define the  $w$ -proportional method as the LC method distributing the first unit according to the above probability (Lemma 3). Equivalently, the  $w$ -proportional method is the UC method distributing the first unit of deficit according to:

$$q_i(x) = \left( \prod_{j \in N \setminus i} w_{ij}^{x_j} \right) \cdot \frac{D(x - e_i)}{D(x)} \quad \text{for all } i \in S(x) \quad (11)$$

Thirdly, under the  $w$ -proportional method the probability distribution of the random variable  $r(t, x)$  is computed as follows:

$$r(t, x)(z) = \left( \prod_{(i,j) \in P} w_{ij}^{z_j(x_i - z_i)} \right) \cdot \frac{D(z) \cdot D(x - z)}{D(x)} \quad (12)$$

**Lemma 6** *The three definitions above coincide. In particular, the  $w$ -proportional method meets LC, UC, PS and PR.*

The proof is omitted for brevity. In some simple cases, the function  $D(x)$  can be computed explicitly. One instance is when  $w_{ij} = 1/2$  for all  $i, j$  where we find the proportional method. Indeed (8) reads in this case:

$$w(s) = 2^{-\sum_{P^*} x_i x_j} \text{ for all } s \in \Delta(x), \text{ and } D(x) = 2^{-\sum_{P^*} x_i x_j} \cdot \frac{x_N!}{\prod_N x_i!}$$

from which the proportional probability follows in (10) and (11). Notice that (12) yields the probability distribution of the proportional method:

$$r(t, x)(z) = \frac{t!(x_N - t)!}{x_N!} \prod_N \frac{x_i!}{z_i!(x_i - z_i)!}$$

The second simple case is for two agents  $N = \{1, 2\}$ . In this case the weighted proportional methods form a one-dimensional family and we set  $\lambda_1 = w_{12}, \lambda_2 = w_{21}$ , with  $0 < \lambda_i < 1$  and  $\lambda_1 + \lambda_2 = 1$ . The function  $D$  can be computed explicitly. Define for all integer  $m, m \geq 1$ :

$$\alpha(m) = \sum_1^m \lambda_1^{m-k} \cdot \lambda_2^{k-1} \text{ and } A(m) = \prod_1^m \alpha(k) \text{ with the convention } \alpha(1) = 1$$

then we have:

$$D(x_1, x_2) = \frac{A(x_1 + x_2)}{A(x_1) \cdot A(x_2)}$$

which follows easily from the recursive property (9) and the limit conditions  $D = 1$  on the axis. Therefore the probabilities  $p$  and  $q$  ((10),(11)) take the form:

$$p_i(x_1, x_2) = \lambda_j^{x_j} \cdot \frac{\alpha(x_i)}{\alpha(x_1 + x_2)}; \quad q_i(x_1, x_2) = \lambda_i^{x_j} \cdot \frac{\alpha(x_i)}{\alpha(x_1 + x_2)} \quad \text{for } \{i, j\} = \{1, 2\} \quad (13)$$

We go back to the general case where  $|N|$  is arbitrary.

**Theorem 3** *The family of  $w$ -proportional methods is characterized by the combination of Lower Composition, Upper Composition and Positive Shares. The statement is preserved if the latter axiom is replaced by Positive Risks.*

Within the family of  $w$ -proportional methods, the Demand Monotonicity and Demand Monotonicity\* requirements impose some restrictions on the weights  $w_{ij}$ . Specifically, the relation  $i \succsim j$  iff  $w_{ij} \geq w_{ji}$  must be transitive<sup>21</sup>; yet we do not know if this condition is sufficient to ensure DM and DM\*.

The entire family of probabilistic rationing methods satisfying Lower and Upper Composition is significantly more complex than its two subsets just discussed, namely the quasipriority (Definition 4) and weighted proportional methods (Lemma 6), except if we have only two agents.

If  $|N| = 2$ , the set of LC+UC methods consists exactly of the two fixed priority methods and the  $w$ -proportional methods: this is Lemma 4 in Moulin [1999a].

For three or more agents there are a large number of methods that satisfy LC and UC but not PS (or PR). One class of examples arises by allowing some of the weights to be zero. A weight  $w_{ij} = 0$  corresponds to asking that agent  $i$  has absolute priority over agent  $j$ . That is  $j$  should not receive any units until  $i$ 's claim is completely filled. If there is no cycle of zero weights, then we still have  $D(x) > 0$  for all  $x$  and we can still define a scheduling method satisfying LC and UC by formulas (8)-(12). This is still the unique scheme with these weights, although the proof of Theorem 3 requires a slight modification.

This leaves the case where there is a cycle of agents, each one having absolute priority over the next. In this case,  $D(x) = 0$  for all  $x$  and the formulas (8)-(12) are meaningless. However for a cycle of zero weights a large number of scheduling methods satisfying LC and UC can be derived from these formulas. Instead of setting certain weights equal to zero, consider taking a limit as those weights approach zero, possibly at different rates. Formulas (10)-(12) are rational functions in the weights and are bounded between 0 and 1 for positive

---

<sup>21</sup>The proof of this claim is available upon request from the authors.

weights. Therefore any sequence of weights tending to zero will have a convergent subsequence and any subsequential limit will be a scheduling method satisfying LC and UC. This leads to a large family of scheduling methods.

However, even this does not exhaust the full scope of LC and UC methods. The limit methods tend to behave much like the quasipriority methods. The units are allocated by a method which is nearly a priority method, however which (nearly) priority method is used depends upon the details of the claims. It is possible to construct more complicated algorithms, where for any given set of claims the scheduling method appears to be a limit of w-proportional methods, but which limit it resembles depends upon the exact claims.

For three agents, this leads to the following infinite parameter family of methods. Suppose the three agents are  $i, j$ , and  $k$  and we want weights  $w_{ij} = 0$ ,  $w_{jk} = 0$ , and  $w_{ki} = 0$ . To completely describe the scheduling method, one chooses a preordering  $\preceq$  of  $\{i, j, k\} \times (N \setminus \{0\})$  which is strictly increasing in the second coordinate. Furthermore for each indifference class  $C$  of the preordering, we choose positive weights  $\omega_i(C)$ ,  $\omega_j(C)$ , and  $\omega_k(C)$ . These weights should sum to 1, and if an agent is not represented in that indifference class his weight should be zero.

The scheduling algorithm is as follows. If only two of the three agents make nonzero claims, then we distribute the units using the priority method given by the zero weight ( $i > j$ ,  $j > k$ , or  $k > i$ ). If all three agents make nonzero claims  $x_i$ ,  $x_j$  and  $x_k$  and one of the pairs  $(i, x_i)$ ,  $(j, x_j)$ , or  $(k, x_k)$  has uniquely the highest rank, then we fill the claims using a priority method  $i > j > k$ ,  $k > i > j$  or  $j > k > i$  chosen so the highest rank claim is filled second. If there is a tie for the highest rank, then the situation is slightly more complicated. Suppose  $(i, x_i)$  and  $(j, x_j)$  are of equal rank and greater rank than  $(k, x_k)$ . Then we will begin by filling part of  $k$ 's claim (possibly none), then fill all of  $i$ 's claim, then all of  $j$ 's claim, then finally the remainder of  $k$ 's claim. The distribution of the amount of  $k$ 's claim filled in the initial segment depends upon the ratio of the weights  $\omega_i(C)$  and  $\omega_j(C)$  for the indifference class containing  $(i, x_i)$  and  $(j, x_j)$ . A similar description holds for a three way tie.

For three agents these are in fact the only possibilities, but analyzing the possibilities for more agents seems infeasible.

## 10 Concluding comment

We have characterized three rich families of probabilistic rationing/scheduling methods by combining, successively CSY and LC (standard of gains: Theorem 2), CSY and UC (standard of losses: Theorem 2), LC and UC (weighted proportional: Theorem 3). What happens when we combine these three powerful properties?

Theorem 2 in Moulin [1999a] answers this question: every such method works by partitioning  $N$  into priority classes  $N_1 \cup N_2 \cup \dots \cup N_K$  — so the agents in  $N_k$  are served in full before those in  $N_{k+1}$  get anything —, using the proportional method within each priority class with three or more agents, and using one of

the LC+UC methods (given by (13)) in each priority class of two agents.

This set of methods is quite inflexible from the standpoint of equity: the only way to depart from the proportional method is to introduce absolute priority between two subgroups of agents. Thus every method but the proportional one is severely inequitable, which stands in sharp contrast with the three families characterized in this paper, where many equitable or nearly equitable methods are available.

## References

- [1] Aumann, R. J. and M. Maschler. 1985. "Game Theoretic Analysis of a Bankruptcy Problem from the Talmud," *Journal of Economic Theory*, 36, 195–213.
- [2] Balinski, M. and H. P. Young. 1982. *Fair Representation: Meeting the Ideal of One Man, One Vote*, Yale University Press, New Haven.
- [3] Banker, R. 1981. "Equity Considerations in Traditional Full Cost Allocation Practices: An Axiomatic Perspective," *Joint Cost Allocations*, S. Moriarty, Ed., Oklahoma City: University of Oklahoma Press, 110–130.
- [4] Barbera, S., M. Jackson and A. Neme. 1997. "Strategyproof Allotment Rules," *Games and Economic Behavior*, 18, 1–21.
- [5] Bazaraa, M., J. Jarvia and H. Sherali. 1990. *Linear Programming in Network Flows*, 2nd Edition. New York, Wiley.
- [6] Demers, A., S. Keshav, and S. Shenker. 1990. "Analysis and Simulation of a Fair Queuing Algorithm," *Internetworking: Research and Experience*, 1, 3–26.
- [7] Elster, J. 1992. *Local Justice*, New York: Russell Sage Foundation.
- [8] Kaminski, M. 2000. "'Hydraulic' rationing," *Mathematical Social Sciences*, 40, 2, 131-156.
- [9] Hofstee, W. 1990. "Allocation by Lot: A Conceptual and Empirical Analysis," *Social Science Information*, 29, 745–763.
- [10] Moulin, H. 1999a. "The Proportional Random Allocation of Indivisible Units", forthcoming *Social Choice and Welfare*.
- [11] Moulin, H. 1999b. "Rationing a commodity along fixed paths," *Journal of Economic Theory*, 84, 41–72.
- [12] Moulin, H. 1999c. "Axiomatic Cost and Surplus Sharing," forthcoming, *Handbook of Social Choice and Welfare*, Arrow, Sen and Suzumura Eds, Elsevier.
- [13] Moulin, H. 2000. "Priority Rules and Other Asymmetric Rationing Methods," *Econometrica*, 68, 3, 643-684.
- [14] O'Neill, B. 1982. "A Problem of Rights Arbitration from the Talmud," *Mathematical Social Sciences*, 2, 345–371.
- [15] Sasaki, H. 1996. "Randomized Uniform Allocation Mechanism and Single-Peaked Preferences of Indivisible Good," mimeo, Waseda University, Japan.



- [16] Shenker, S. 1995, circulated 1989. “Making greed work in networks: A game-theoretic analysis of gateway service disciplines,” *IEEE/ACM Transactions on Networking*, 3, 6, 819–31.
- [17] Thomson, W. 1995. “Axiomatic Analyses of Bankruptcy and Taxation Problems: A Survey,” mimeo, University of Rochester.
- [18] Young, H. P. 1988. “Distributive Justice in Taxation,” *Journal of Economic Theory*, 48, 331–335.
- [19] Young, H. P. 1994. *Equity: in Theory and Practice*, Princeton University Press, Princeton.

# Appendix: Proofs

## 1. Proposition 1

It is possible to deduce the result from Theorem 2. However, the direct proof of Proposition 1 is very simple and serves to support the intuition for the more involved proof of Theorem 2.

### a) Statement “only if”

A s.g. method meets LC by definition. To check DM, we fix a profile  $x$  and an agent  $i$ , and compare the sequences  $r(x)$  and  $r(x + e_i)$ . If  $i$  appears at times  $t_1, t_2, \dots, t_{x_i}$  in  $r(x)$ , then  $i$  appears exactly once in the first  $t_1$  terms of  $r(x + e_i)$ , and is the  $(t_1 + 1)$ -th term of that sequence; moreover the last  $(x_N - t_1)$  terms of  $r(x)$  and  $r(x + e_i)$  coincide. DM follows.

Consistency means that the  $(i, j)$ -subsequence of  $r(x)$  (by deleting all agents other than  $i$  and  $j$ ) is independent of  $x_k$  for  $k \neq i, j$ . But this subsequence obtains from ordering the pairs  $(i, y_i), 1 \leq y_i \leq x_i$ , and  $(j, y_j), 1 \leq y_j \leq x_j$  according to  $\succ$ .

### b) Statement if

Let  $r$  be a method meeting CSY, LC and DM. We write  $\Pi(x) = i$  iff  $r(1, x) = e_i$ . We define a binary relation  $\succ$  on  $N \times (\mathbb{N} \setminus \{0\})$ :

$$(i, x_i) \succ (i, x'_i) \text{ iff } x_i > x'_i, \text{ and for } i \neq j : (i, x_i) \succ (j, x_j) \text{ iff } \Pi(x_i e_i + x_j e_j) = i$$

The relation  $\succ$  is clearly complete and antisymmetric; we check its transitivity:

$$\begin{aligned} (i, x_i) \succ (i, x'_i) \succ (j, x_j) &\Rightarrow (i, x_i) \succ (j, x_j) \text{ by DM} \\ (i, x_i) \succ (j, x_j) \succ (j, x'_j) &\Rightarrow (i, x_i) \succ (j, x'_j) \text{ by DM} \\ (i, x_i) \succ (j, x_j) \succ (i, x'_i) &\Rightarrow x_i > x'_i \text{ by the above property} \end{aligned}$$

It remains to check the case of three different agents  $i, j, k$  such that  $(i, x_i) \succ (j, x_j)$  and  $(j, x_j) \succ (k, x_k)$ . Denote  $x = x_i e_i + x_j e_j + x_k e_k$  and consider the sequence  $r(x)$ . It cannot start by  $j$  otherwise CSY would imply that  $r(x_i e_i + x_j e_j)$  starts by  $j$  as well. It cannot similarly start by  $k$ , so  $r(x)$  starts by  $i$ . Then CSY implies that  $r(x_i e_i + x_k e_k)$  starts by  $i$  as was to be shown.

We pick now an arbitrary profile  $x$  and let  $(i, x_i)$  be the unique maximum of  $\succ$ . By  $(i, x_i) \succ (j, x_j)$  and CSY,  $r(x)$  cannot start with  $j$ ; as this is true for all  $j, j \neq i$ , we conclude  $\Pi(x) = i$ . By Lemma 1,  $r$  is the s.g. method associated with  $\succ$ .

## 2. Theorem 1, statements b and c

a) The first definition of fair queuing makes clear that it satisfies DM; the dual definition shows UC; the two equity properties are obvious.

b) We prove that a method meeting LC, DM, ETEA, and ETEP must be fair queuing\*. Our method is described by the function  $p(x)$  as in Lemma 3.

We use the following notation for all  $x$  in  $\mathbb{N}^N$  :

$$\bar{x} = \max_i x_i, T(x) = \{i \mid x_i = \bar{x}\}, R(x) = \{i \mid x_i = \bar{x} - 1\}$$

and speak of a top agent in  $T(x)$  and a runner up in  $R(x)$ . We denote by  $A(t, t')$  the set of profiles  $x$  with  $t$  top agents and  $t'$  runner-ups:  $|T(x)| = t$  and  $|R(x)| = t'$ . Thus  $t \geq 1, t' \geq 0$  and  $t + t' \leq n$ .

We must show that  $p(x)$  is the uniform distribution over  $T(x)$ . We prove it on  $A(t, t')$  by induction on  $t$  and  $t'$ . Write  $P(t, t')$  if  $p$  is as desired on  $A(t, t')$ .

$P(n, 0)$  follows from ETEA. Assume  $P(t, n - t)$  holds for some  $t \geq 2$  and show  $P(t - 1, n - t + 1)$ . Choose  $x$  in  $A(t - 1, n - t + 1)$  and  $i$  a runner up in  $x$ . As  $x + e_i$  is in  $A(t, n - t)$  we have  $p_i(x + e_i) > 0$ . Suppose  $p_i(x) > 0$  : then when we allocate two units at profile  $x + e_i$ , agent  $i$  might get both units:

$$\text{proba}\{r_i(2, x + e_i) = 2\} = p_i(x + e_i) \cdot p_i(x) > 0$$

This contradicts ETEP, because in  $x + e_i$  agent  $i$  and any  $j$  in  $T(x)$  have equal demands. Hence  $p_i(x) = 0$  for any  $i$  in  $R(x)$ . As  $T(x) = N \setminus R(x)$ , ETEA implies that  $p$  is uniform on  $T(x)$ .

We have shown  $P(t, n - t)$  for all  $t = 1, \dots, n$ . Next we show  $P(t, t')$  for all  $t, t'$  by *decreasing* induction on  $t + t'$ . Let  $x$  be an arbitrary profile in  $A(t, t')$  and  $i$  be a runner up in  $x$ . As  $x + e_i$  is in  $A(t + 1, t')$ , the inductive assumption shows  $p_i(x + e_i) > 0$ . As in the previous paragraph,  $p_i(x) > 0$  would bring a contradiction of ETEP. Thus  $p_i(x) = 0$  for all runner-ups. Next we pick  $i$  outside of  $T(x)$  and  $R(x)$ , namely  $x_i \leq \bar{x} - 2$ . Consider  $x' = x + (\bar{x} - x_i - 1) \cdot e_i$ , and note that  $i$  is a runner-up in  $x'$  and  $x' \in A(t, t' + 1)$ . By induction  $p_i(x') = 0$ , hence by DM  $p_i(x) = 0$  as well. We have shown that the support of  $p(x)$  is  $T(x)$  and the desired conclusion follows from ETEA.

### 3. Lemma 4

The “only if” statement is clear. To prove “if”, we fix  $z$  and a function  $p^z$  as in the statement of Lemma 4 and we show that the corresponding method  $r^z$  satisfies DM. As  $z$  remains fixed throughout the proof we simply write  $p_i(x_i, x_j)$  for the probability that  $i$  receives the first unit given claims  $x_i e_i + x_j e_j$ . We use the following notation:

$$P(a, t; x_i, x_j) = \text{proba. } i \text{ gets at most } a \text{ units under } r \text{ at problem } (t; x_i, x_j)$$

We must show  $P(a, t; x_i, x_j) \geq P(a, t; x_i + 1, x_j)$ . To this end we compute:

$$\begin{aligned} & P(a, t + 1; x_i, x_j) \\ = & \text{proba}(i \text{ gets } a \text{ of first } t \text{ and } j \text{ gets } (t + 1) - st) + \text{proba}(i \text{ gets } \leq a - 1 \text{ of first } t) \\ = & [P(a, t; x_i, x_j) - P(a - 1, t; x_i, x_j)] \cdot p_j(x_i - a, x_j + a - t) + P(a - 1, t; x_i, x_j) \\ = & P(a, t; x_i, x_j) \cdot p_j(x_i - a, x_j + a - t) + P(a - 1, t; x_i, x_j) \cdot p_i(x_i - a, x_j + a - t) \end{aligned}$$

Thus,

$$\begin{aligned}
& P(a, t + 1; x_i, x_j) - P(a, t + 1; x_i + 1, x_j) \\
= & [P(a, t; x_i, x_j) - P(a, t; x_i + 1, x_j)] \cdot p_j(x_i - a, x_j + a - t) \\
& + [P(a - 1, t; x_i, x_j) - P(a - 1, t; x_i + 1, x_j)] \cdot p_i(x_i - a, x_j + a - t) \\
& + [P(a, t; x_i + 1, x_j) - P(a - 1, t; x_i + 1, x_j)] \cdot \\
& \quad (p_i(x_i + 1 - a, x_j + a - t) - p_i(x_i - a, x_j + a - t))
\end{aligned}$$

The last term is non negative by assumption. The previous two are non negative if we assume that the desired inequality holds for  $t$ . Thus we conclude by induction on  $t$ .

#### 4. **Theorem 2**

a) A standard of gains method meets CSY and DM. Fix a method  $r$  as in Definition 3. We use a representation of  $r(x)$  similar to that in Remark 4. Given a demand profile  $x$ , there are finitely many indifference classes intersecting with  $\cup_N(i, x'_i)$ ,  $1 \leq x'_i \leq x_i$ , and we can order them increasingly as  $\delta^1, \dots, \delta^Q$ , with corresponding depth  $z^q = \theta(\delta^q)$ . Then we set:

$$x^q = x \wedge (z^1 + \dots + z^q) = (x - x^{q-1}) \wedge z^q + x^{q-1} \text{ for } q = 1, \dots, Q \text{ (with } x^0 = 0)$$

in particular  $x^Q = x$ . For any  $t, 1 \leq t \leq x_N - 1$ , we find the unique index  $q^*$  such that  $x_N^{q^*-1} \leq t < x_N^{q^*}$  and the random variable  $r(t, x)$  can now be written:

$$r(t, x) = x^{q^*-1} + r^{z^{q^*}}(t - x_N^{q^*-1}, x^{q^*} - x^{q^*-1}) \quad (14)$$

in which the randomness only comes from the allocation of the last  $t - x_N^{q^*-1}$  units.

To check CSY, we “drop” agent 1 and compute the allocation of  $t - r_1(t, x)$  units at profile  $(x \mid^1 0)$ . By (14) any realization  $y_1$  of the random variable  $r_1(t, x)$  is s.t.  $x_1^{q^*-1} \leq y_1 \leq x_1^{q^*-1} + t - x_N^{q^*-1} = t - x_{N \setminus 1}^{q^*-1}$ . Moreover, setting  $y'_1 = y_1 - x_1^{q^*-1}$ , (14) also implies  $t - x_N^{q^*-1} - y'_1 \leq x_{N \setminus 1}^{q^*} - x_{N \setminus 1}^{q^*-1}$ . Rearranging these inequalities, we get:

$$x_{N \setminus 1}^{q^*-1} \leq t - y_1 \leq x_{N \setminus 1}^{q^*} \quad (15)$$

On the other hand the sequence of indifference classes associated with  $\tilde{x} = (x \mid^1 0)$  is the subsequence of  $\delta^1, \dots, \delta^Q$  where we drop  $\delta^q$  if it only intersects with pairs like  $(1, x'_1)$ . Thus the corresponding sequence  $\tilde{x}^q$  is simply  $\tilde{x}^q = (x^q \mid^1 0)$ , provided we allow a repetition  $\tilde{x}^q = \tilde{x}^{q-1}$  for each dropped class. Inequalities

(15) mean that the critical index  $q^*$  is the same for  $(t, x)$  and for  $(t - y_1, \tilde{x})$ , for any realization  $y_1$ . Therefore:

$$\begin{aligned} r(t - y_1, (x \mid^1 0)) &= \tilde{x}^{q^*-1} + r^{z^{q^*}} (t - y_1 - \tilde{x}_N^{q^*-1}, \tilde{x}^{q^*} - \tilde{x}^{q^*-1}) \\ r(t - y_1, (x \mid^1 0)) &= \tilde{x}^{q^*-1} + r^{z^{q^*}} (t - \tilde{x}_N^{q^*-1} - y'_1, (x^{q^*}, x^{q^*-1} \mid^1 0)) \\ \text{where } y'_1 &= y_1 - x_1^{q^*-1} \end{aligned}$$

Comparing the above equality with (14), it is now clear that the consistency of  $r$  follows from that of  $r^z$ .

To check DM, we note that increasing  $x$  to  $x + e_i$  can only add one indifference class  $\delta^{Q+1}$  to the sequence  $\delta^q$ . If no class is added, then the critical index  $q^*$  is the same at  $(t, x)$  and  $(t, x + e_i)$ , and two cases may arise. If  $(x + e_i)^{q^*-1} = x^{q^*-1} + e_i$  then  $(x + e_i)_i^{q^*} - (x + e_i)_i^{q^*-1} = 0$  and the share of  $i$  increases by one. If  $(x + e_i)^{q^*-1} = x^{q^*-1}$  then the desired inequality follows from the DM property of  $r^z$ .

b) A method meeting CSY, LC and DM is a standard of gains

We fix a LC method  $r$  associated with the function  $x \rightarrow p(x)$  (*Lemma 3*) and assume the method is DM and CSY. We must show that  $p$  takes the form described in Definition 3.

Step 1 – Constructing the standard of gains  $\succsim$

We define first a complete binary relation  $R$  on  $N \times (N \setminus \{0\})$ :

$$(i, x_i) R (j, x_j) \stackrel{def}{\Leftrightarrow} \{i = j \text{ and } x_i \geq x_j\} \text{ or } \{i \neq j \text{ and } p_i(x_i e_i + x_j e_j) > 0\}$$

Although  $R$  may not be transitive (as explained below), it is almost transitive, in the following sense:

$$\begin{aligned} \text{for } i \neq j : (i, x_i) R (j, x_j) \text{ and } (j, x_j) R (j, x'_j) &\Rightarrow (i, x_i) R (j, x'_j) \\ \text{for } i \neq j : (i, x_i) R (i, x'_i) \text{ and } (i, x'_i) R (j, x_j) &\Rightarrow (i, x_i) R (j, x_j) \\ \text{for } i \neq j \neq k \neq i : (i, x_i) R (j, x_j) \text{ and } (j, x_j) R (k, x_k) &\Rightarrow (i, x_i) R (k, x_k) \end{aligned} \tag{16}$$

To prove the top statement observe:

$$\begin{aligned} (i, x_i) R (j, x_j) &\Leftrightarrow 1 > p_j(x_i e_i + x_j e_j) \\ \text{and } x_j &\geq x'_j \Rightarrow p_j(x_i e_i + x'_j e_j) \leq p_j(x_i e_i + x_j e_j) \end{aligned}$$

where the second implication follows from DM. The proof of the middle statement is similar. To prove the last statement we set  $x = x_i e_i + x_j e_j + x_k e_k$  and proceed by contradiction: we assume  $p_i(x_i e_i + x_k e_k) = 0$  (the negation of  $(i, x_i) R (k, x_k)$ ),  $p_i(x_i e_i + x_j e_j) > 0$  and  $p_j(x_j e_j + x_k e_k) > 0$ . Apply CSY at  $t = 1$ :  $p_i(x) = (1 - p_j(x)) \cdot p_i(x_i e_i + x_k e_k)$ . Therefore  $p_i(x_i e_i + x_k e_k) = 0$  implies

$p_i(x) = 0$ . By CSY again:  $p_i(x) = (1 - p_k(x)) \cdot p_i(x_i e_i + x_j e_j)$ , hence  $p_k(x) = 1$  and  $p_j(x) = 0$ . Apply CSY one more time  $p_j(x) = (1 - p_i(x)) \cdot p_j(x_j e_j + x_k e_k)$  to reach the desired contradiction.

Properties (16) do not yield the transitivity of  $R$ , because chains alternating between two agents are possible:

$$(i, x_i) R (j, x_j) R (i, x'_i) R (j, x'_j) R \dots \text{ with } x_i < x'_i < \dots \text{ and } x_j < x'_j < \dots \quad (17)$$

But (16) implies that the above are the only inclusion minimal  $R$  chains.

We define the preordering  $\succsim$  as the transitive closure of  $R$ . Thus  $(i, x_i + 1) \succsim (i, x_i)$  and  $\succsim$  is a standard of gains: Definition 3. In the rest of the proof, we use the notations introduced after Definition 3 for the indifference classes, their support, depth, etc...

Step 2 – Properties linking  $\succsim$  and  $p(x)$

Denote by  $\succ$  the strict component of  $\succsim$ . We note:

$$(i, x_i) \succ (j, x_j) \Rightarrow p_i(x_i e_i + x_j e_j) = 1$$

Indeed, if  $p_j(x_i e_i + x_j e_j) > 0$ , we have  $(j, x_j) \succsim (i, x_i)$ . Next we consider an arbitrary profile  $x$  and denote by  $T(x)$  the set of agents whose claims are in the highest indifference class of  $\succsim$ :  $i \in T(x)$  iff  $(i, x_i) \succsim (j, x_j)$  for all  $j$ . Our first property is:

$$p_j(x) = 0 \text{ for } j \notin T(x) \text{ and } p_i(x) = p_i(x \upharpoonright^{N \setminus T(x)} 0) \text{ for } i \in T(x) \quad (18)$$

We fix  $x$ , two arbitrary agents denoted 1,2 and we assume  $p_1(x) > 0$ . For all  $i \geq 3$ , CSY implies:

$$p_1(x) = (1 - p_i(x)) \cdot p_1(x \upharpoonright^i 0) ; p_2(x) = (1 - p_i(x)) \cdot p_2(x \upharpoonright^i 0)$$

The left hand equality yields  $p_1(x \upharpoonright^i 0) > 0$  hence we have:

$$\frac{p_2(x)}{p_1(x)} = \frac{p_2(x \upharpoonright^i 0)}{p_1(x \upharpoonright^i 0)}$$

Repeating this argument, we successively eliminate all agents  $i \geq 3$ :

$$\frac{p_2(x)}{p_1(x)} = \frac{p_2(x_1 e_1 + x_2 e_2)}{p_1(x_1 e_1 + x_2 e_2)} \quad (19)$$

In particular  $p_1(x) > 0 \Rightarrow p_1(x_1 e_1 + x_2 e_2) > 0 \Rightarrow (1, x_1) \succsim (2, x_2)$ .

As the choice of agent 2 was arbitrary, we conclude  $p_1(x) > 0 \Rightarrow 1 \in T(x)$ . To prove the second statement in (18), we pick  $1 \notin T(x)$  and invoke CSY:

$$p_i(x) = (1 - p_1(x)) \cdot p_i(x \mid^1 0) = p_i(x \mid^1 0) \text{ for all } i \geq 2$$

Repeat this argument to eliminate successively all agents in  $N \setminus T(x)$ .

Our next property applies only to an indifference class of type three or more:

$$\{\text{type}(\delta) \geq 3 \text{ and } (i, x_i), (j, x_j) \in \delta\} \Rightarrow 0 < p_i(x_i e_i + x_j e_j) < 1 \quad (20)$$

Choose a third agent  $k$  in the support of  $\delta$ , and  $(k, x_k) \in \delta$ . We have  $(i, x_i) \succsim (k, x_k)$  and  $(k, x_k) \succsim (j, x_j)$ . We know from Step 1 that an inclusion minimal R-chain from  $(i, x_i)$  to  $(k, x_k)$  or from  $(k, x_k)$  to  $(j, x_j)$  takes the form (17). Linking two such chains and using (16) repeatedly, we find  $(i, x_i) R (j, x_j)$ ! Exchanging the roles of  $i$  and  $j$  gives similarly  $(j, x_j) R (i, x_i)$  and (20) is established.

The last property applies when the highest indifference class among  $(i, x_i)$ , denoted  $\delta(x)$ , is of type 3 or more. For any demand profile  $x$ , we have:

$$\text{if type}(\delta(x)) \geq 3 : p_i(x) > 0 \Leftrightarrow i \in T(x) \quad (21)$$

We already know that the support of  $p(x)$  is contained in  $T(x)$  (property (18)). To show the converse inclusion, pick two agents 1 and 2 in  $T(x)$ , with agent 1 in the support of  $p(x)$  (if no two such agents can be found,  $T(x)$  is a singleton and there is nothing to prove). As above,  $p_1(x) > 0$  implies equation (19). Because  $\delta(x)$  contains  $(1, x_1)$  and  $(2, x_2)$ , the right hand ratio is positive (by property (20)), hence  $p_2(x)$  is positive as desired.

#### Step 3 – End of proof

On an indifference class  $\delta$  of  $\succsim$  with support  $S$  and depth  $z = \theta(\delta)$ , the probability function  $p$  induces a LC method restricted to  $[0, z]$  as follows:

for all  $y \in [0, z] : p^z(y) = p(y + \alpha)$  where  $\alpha_i = a_i - 1$  for  $i \in S$ ,  $= 0$  otherwise

(here  $I_i = [a_i, b_i]$  as in Definition 2).

We check that the probability functions  $p^z$ , one for each  $\delta$ , determine  $p$  entirely in the following sense: for all  $x : p(x) = p^z(d(x))$  where  $p^z$  corresponds to  $\delta(x)$ .

Fix  $x$  and recall that  $T(x)$  is a subset of  $S$ , the support of  $\delta(x)$ . Hence for all  $i \in T(x)$  we have:  $(d(x) + \alpha)_i = (x_i - a_i + 1) + (a_i - 1) = x_i$ . Thus  $x$  and  $d(x) + \alpha$  coincide on  $T(x)$ ; moreover  $T(d(x) + \alpha) = T(x)$  by the definition of  $d(x)$  and the fact that for all  $i$  in  $S$ , the support of  $\delta(x)$ , we have  $\delta(x) \succ (i, a_i - 1)$ . These two observations combined with property (18) imply:  $p(x) = p(d(x) + \alpha) = p^z(d(x))$  as desired.

Next we show that for any indifference class  $\delta$ , the corresponding method  $r^z$  on  $[0, z]$  meets DM and CSY. Observe that:

$$r^z(t, y) = r(t, y + \alpha) \text{ for } 0 \leq y \leq z, t \leq y_N$$

This follows from  $p^z(y') = p(y' + \alpha)$  for  $0 \leq y' \leq y$  by repeated applications of LC. Now  $r^z$  is DM because  $r$  is. As for CSY, we have:

$$r_i^z(t - r_j^z(t, y), (y \mid^j 0)) = r_i(t - r_j(t, y + \alpha), (y \mid^j 0) + \alpha)$$

We note that at profile  $x = (y \mid^j 0) + \alpha$ , agent  $j$  gets none of the first  $y_{N \setminus j}$  units because  $\delta(x) \succ (j, a_j - 1)$  and it takes  $y_{N \setminus j}$  units to exhaust the class  $\delta(x)$ . Thus we can (by CSY of  $r$ ) replace  $(y \mid^j 0) + \alpha$  in the righthand term by  $(y + \alpha \mid^j 0)$  and the desired conclusion follows, again by CSY.

The only thing left to prove is that over an indifference class  $\delta$  of type 3 or more,  $p^z$  is the proportional method. First we note that  $p^z$  satisfies the Positive Share property

$$\text{for all } y, 0 \leq y \leq z : y_i > 0 \Rightarrow p_i^z(y) > 0$$

Indeed, set  $x = y + \alpha$  and note that  $T(x)$  coincides with the support of  $y$ . By (21), the support of  $p(x) = p^z(y)$  is  $T(x)$  as well.

We conclude with the help of Theorem 3 in Moulin [1999a], establishing that with three agents or more, the combination CSY, LC and PS is met only by the proportional method. It is a simple matter, omitted for brevity, to adapt this proof to the case of restricted rationing methods.

## 5. Proposition 2

Suppose we have a claim of  $x$  which is distributed by a method that satisfies LC and UC. We consider the sequence  $(j_1, j_2, \dots, j_{x_N})$  in which the claims are to be distributed. Note that the combination LC and UC implies that if we select any substring of this sequence which distributes a total claim of  $y$ , then it is in fact exactly the order in which we would fill claims of  $y$ .

Suppose we fill part of agent  $i$ 's claim, then fill parts of other claims, then return to agent  $i$ 's claim. Then part of our sequence will read  $(i, j_k, \dots, i)$ . Pulling the first element off this subsequence gives a string which fills the same total claim  $y$  as pulling the last element off this subsequence. By LC and UC, these must both be the unique sequence corresponding to claims of  $y$ . This is a contradiction. Thus we see that each agent's claim must be completely filled in a single block. Rephrasing this, we see that any deterministic method satisfying LC and UC must have the property that claims are always filled by a priority method, though which of the  $|N|!$  priority methods is used may depend upon the claims.



Suppose we increase the claim of agent  $i$ . DM says that we cannot increase the number of units allotted before we fill agent  $i$ 's claim and DM\* says that we cannot increase the number of units allotted after agent  $i$ 's claim. Thus the agents served before agent  $i$  may change, but the sum of their claims may not change. Notice that LC and UC imply that if the agents served before agent  $i$  do not change, then the priority method used cannot change.

Fix a total claim  $x$  with nonzero claims by a set of agents  $M$ . Let  $z$  be a new claim, with nonzero claims by the same set of agents  $M$ , but with each such agent making a claim of  $v > x_M$ . We will show that the claims  $x$  and  $z$  are filled using the same priority method. From this it follows that the priority method depends only on the agents making nonzero claims. Suppose for the claim of  $x$ , the priority method used is  $1 > 2 > \dots > m$ .

As we increase agent 1's claim to  $v$ , DM implies that he must remain the first agent served and LC implies that the order for the other agents is unchanged. Thus the priority method used does not change. Having already increased agent 1's claim to  $v$ , suppose we increase agent 2's claim to  $v$ . As remarked above DM and DM\* imply that the number of units distributed before agent 2's claim is filled cannot change. Since  $v > x_M$ , we see that agent 1 must still be the only agent served before agent 2. Hence the priority method used cannot change. Continuing in this way, we see that as we move from  $x$  to  $z$  increasing the claims of agents in order, the priority method used can never change. Thus  $x$  and  $z$  are filled using the same priority. Hence the priority method used depends only on the agents making nonzero claims.

Finally note that for any rationing/scheduling method which selects for each set  $S$  of nonzero claims an ordering  $\sigma_S$  of  $S$  and uses the corresponding priority method, the conditions

$$\begin{aligned} \text{top}(\sigma_S) &= i \Rightarrow \sigma_{S \setminus i} \text{ and } \sigma_S \text{ coincide on } S \setminus i \\ \text{bot}(\sigma_S) &= j \Rightarrow \sigma_{S \setminus j} \text{ and } \sigma_S \text{ coincide on } S \setminus j \end{aligned}$$

are equivalent to LC and UC, respectively. Thus we have a quasipriority method.

Conversely, this shows that any quasipriority satisfies LC and UC. Also from the discussion above, we see that since the order of the claims is unchanged by increasing a nonzero claim, DM and DM\* hold in this case. If a claim is increased from 0 to 1, then DM and DM\* are vacuous. Thus any quasipriority satisfies DM and DM\*.

### 6. Theorem 3

Suppose we have an allocation method which satisfies LC, UC and PS. Let  $p_j(x)$  denote the probability that agent  $j$  receives the first unit given claims of  $x$  and dually let  $\bar{p}_j(x)$  denote the probability that agent  $j$  receives the first unit of deficit given claims of  $x$ . Define weights  $w_{ij}$  by:

$$w_{ij} = p_j(e_i + e_j) = \bar{p}_i(e_i + e_j)$$

Note that by PS,  $w_{ij} > 0$  and  $w_{ij} + w_{ji} = p_j(e_i + e_j) + p_i(e_i + e_j) = 1$ . We will show by induction on  $x_N$  that this method must agree with the w-proportional method with these weights. Note that they agree if  $x_N \leq 2$  by construction.

Applying LC to  $r(2, x)$ , we have:

$$\begin{aligned} r(1, x)(e_i) &= p_i(x) \\ r(2, x)(2e_i) &= p_i(x)p_i(x - e_i) \\ r(2, x)(e_i + e_j) &= p_i(x)p_j(x - e_i) + p_j(x)p_i(x - e_j) \end{aligned}$$

A consequence of UC is  $r(1, x) = r^2(1, r^1(2, x))$ , hence we get:

$$r(1, x)(e_i) = r(2, x)(2e_i) + \sum_{j \in N \setminus i} r(2, x)(e_i + e_j) \bar{p}_j(e_i + e_j)$$

Plugging in the formulas for  $r$  and  $\bar{p}$  in terms of  $p$  obtained above gives:

$$p_i(x) = p_i(x)p_i(x - e_i) + \sum (p_i(x)p_j(x - e_i) + p_j(x)p_i(x - e_j))p_i(e_i + e_j)$$

Since  $\sum_{j \in N} p_j(x - e_i) = 1$  we can rewrite this as:

$$\left( \sum_{j \in N \setminus i} p_j(x - e_i)p_j(e_i + e_j) \right) p_i(x) = \sum_{j \in N \setminus i} (p_i(x - e_j)p_i(e_i + e_j))p_j(x)$$

Assuming  $p_k(y)$  to be known for  $y_N < x_N$ , we can regard these as  $|N|$  linear equations in the  $|N|$  unknowns  $p_k(x)$ . Note that in matrix form, this gives a system of equations  $Av = 0$ , where (by PS)  $A$  will be an  $|N| \times |N|$  matrix with positive entries off the diagonal and negative entries on the diagonal. The coefficients in  $A$  depend only on the  $p_k(y)$  for  $y_N < x_N$  hence by induction they agree with the coefficients we would get from the w-proportional method. Thus the system  $Av = 0$  has a solution  $u$  with all positive entries corresponding to the  $p_k(x)$  for the w-proportional method. From the Lemma below it follows that the only solutions are multiples of  $u$ . However we have the further equation  $\sum_{k \in N} p_k(x) = 1$ . Thus we see that the w-proportional method gives the unique set of probabilities satisfying these equations. Hence our method agrees with the w-proportional method for a claim of  $x$  and the induction proof is complete.

### 7. Lemma.

Suppose  $A$  is an  $n \times n$  matrix with positive entries off the diagonal and negative entries on the diagonal and suppose the system  $Aw = 0$  has a solution  $u$  with all entries positive. Then the only solutions to  $Aw = 0$  are multiples of  $u$ .

Proof. Suppose  $Aw = 0$  and choose an index  $i$  such that  $|w_i/u_i|$  is maximal. Then the  $i$ th equation in the system gives:

$$(-A_{ii})w_i = \sum_{j \neq i} A_{ij}w_j$$

and similarly for  $u$ . Hence,

$$|A_{ii}w_i| = \left| \sum_{j \neq i} A_{ij}w_j \right| \leq \sum_{j \neq i} A_{ij} |w_j| \leq \sum_{j \neq i} A_{ij}u_j |w_i/u_i| = (-A_{ii})u_i |w_i/u_i| = |A_{ii}w_i|$$

Here the first inequality is the triangle inequality, the second is maximality of  $|w_i/u_i|$  and the last equality follows since  $u_i > 0$  and  $A_{ii} < 0$ . Thus both inequalities must actually be equalities. Thus all the  $w_j$  have the same sign and  $|w_i/u_i|$  is a constant, i.e.  $w$  is a multiple of  $u$ .