Can competition ever be fair?
Challenging the standard prejudice [#]

Christian Arnsperger

FNRS (Brussels), IRES (UCL)
& Chaire Hoover (UCL)
arnsperger@etes.ucl.ac.be

Philippe De Villé

IRES (UCL)
deville@ires.ucl.ac.be

May 2002

Abstract

In this paper, we challenge the usual argument which says that competition is a fair mechanism because it ranks individuals according to their relative preferences between effort and leisure. This argument, we claim, is very insufficient as a justification of fairness in competition, and we show that it does not stand up to scrutiny once various dynamic aspects of competition are taken into account. Once the sequential unfolding of competition is taken into account, competition turns out to be unfair *even if* the usual fairness argument is upheld. We distinguish between two notions of fairness, which we call U-fairness, where "U" stands for the "usual" fairness notion, and S-fairness, where "S" stands for the "sequential" aspect of competition. The sequential unfairness of competition, we argue, comprises two usually neglected aspects connected with losses of freedom: first of all, there is an "eclipse" of preferences in the sense that even perfectly calculating competitors do not carry out a trade-off between effort and ranking; and second, competitive dynamics leads to single-mindedness because the constraints on the competitors' choices *always* operate in the sense of *increased* competitiveness and, therefore, in the direction of an *increased* effort requirements. We argue (1) that competition is S-unfair even if it is U-fair, (2) that as S-unfairness increases, the ethical relevance of U-fairness itself vanishes, so that (3) by focusing as they usually do on U-fairness alone, economists neglect much deeper aspects of unfairness.

JEL Classification Numbers: A13, B41, D50, D63, P10

[#] Financial support from the Belgian French Community's program Action de Recherches Concertée 99/04-235 (IRES, UCL) is gratefully acknowledged. Part of De Villé's contribution is, in addition, based upon work done as Visiting Fellow at the Hoover Institution, Stanford University, whose support is also gratefully acknowledged. Seminar participants at the Hoover Institution (April 1999), Chaire Hoover (December 1999), and IRES (December 2000) offered many stimulating suggestions and objections. We are particularly grateful to Jacques H. Drèze for his detailed and insightful remarks on an earlier version; he should not be held responsible for any errors or inconsistencies.

Introduction

Can competition ever be fair? This question seems to fly in the face of self-evidence among most economists. Their immediate reply might well be: What, if anything, can you suggest that does *better*, in terms of fairness, than a competitive system? Or, worse still, our question is apt to be received with a smile. To many indeed, competition—at least in its "perfect" version of a process so thoroughly purified of any distortions that it can only be virtual—is the paragon of a fair mechanism: *if* you equalise all features which make people initially unequal and for which they cannot be held responsible, *then* the outcome of a non-distorted competition is necessarily fair. In fact, the procedure itself becomes useless; just provide each agent with the equilibrium prices and they will directly "hop" to the outcome. (The result that a competitive equilibrium from equal split of initial resources is envy-free reflects this core intuition.) So why ask the question? Isn't the answer totally obvious?

Our intuition when starting reflecting on the fairness of competition as a principle of socio-economic organisation was that this "obvious" character is, at least partly, rhetorical. We want to show that there is a very powerful presumption against the fairness of competition being a sound theoretical argument, let alone a sound social and political one. The standard fairness-of-competition argument indeed rests crucially on the classic split between resources (both external and internal) and preferences: if you have been given the same resources as everyone else, the outcome of competition will be determined solely by the interaction of your preference-based choices and the other agents' preference-based choices. Thus, it is necessarily a fair outcome, or so the usual argument goes. Our main point in this paper is that this usual argument misses out on crucial aspects of competition as a dynamically unfolding process; we will show that once this dynamic unfolding is taken into account, competition is unfair *even if* the usual fairness argument is upheld. In that sense, our objective is that by defending competition on the basis of the usual fairness argument, one is "barking up the wrong tree": there is something much more *deeply unfair* in competition, something which occurs even if the usual fairness argument is satisfied. This "something" has to do with agents not being able to actually apply their desired trade-off between leisure and effort because the dynamics of competition has progressively restricted their choice sets in a way which makes this trade-off impossible. To put it bluntly, an equal competition at a given point in time, which is deemed fair by the defenders of the usual argument, might in fact be unfair within the broader evaluative framework we advocate here. Thus, we argue that the usual fairness-of-competition argument is not sufficient to be able to judge the fairness of a competitive process: that can only be assessed by looking at *the whole sequence of steps* through which the process has shaped the individual competitors' actual choice sets.

Our discussion will proceed in three parts. In section 1, we set out what we believe to be the view of fair competition which underlies most usual defences of competition as a fair and efficient procedure. Let us call this *usual* fairness criterion "*U-fairness*". This is the usual view we want to challenge in sections 2 and 3. In section 2, we argue that the fundamental dynamics of competition leads to the progressive, sequential curtailement of competitors' choice sets, even if the competition is U-fair at every stage in the process. In section 3, we summarise the ethical issues which we deem crucial and we argue that this process of *sequential* shrinking in the competitors' options creates what we can call *S-unfairness*—an unfairness which arises independently of whether U-fairness is satisfied or not. In fact it will turn out, in the course of our argument, that the more S-unfair the competition becomes, the more U-fairness itself becomes practically irrelevant: the only way for U-fairness to be upheld will be for competition to be completely non-selective, thus losing its hallowed feature as an efficient screening process. Thus, in our view, S-fairness should stand tall alongside U-fairness in the debate on the fairness of competition. Therefore, section 3 challenges the relevance of the concept of a "competitive equilibrium" as an adequate benchmark for the understanding of what the overall fairness-of-competition debate should be about.

1. A formal model of competition: the criterion of U-fairness

Let us first give the usual fairness-of-competition argument its best shot. We want to investigate the precise scope of that argument by setting it within the metaphorical framework of a *race*, in which a fixed number of runners line up, run at the "go" signal, and cross the finish line. Although highly stylised, this framework serves as the implicit background for many discourses on competition being a fair procedure. The fairness-of-competition argument, when transposed to the specific case of a race, says the following: if we have equalised all factors which objectively determine the individual runners' *absolute* (rather than relative) performances, the order of arrival will be completely determined by the various runners' preferences for effort relative to gain, that is, by their *conscious choices*. This can be modelled as follows. Let π_i be the measure of runner *i*'s performance, which in this case is simply the inverse of "running time." (We take the inverse so that the "performance" variable will have the intuitive property of being numerically higher when running performance is better.) The technology for producing performance is simple: the runner needs a certain vector of *m* resources, which we denote

$\mathbf{r}_i \quad R^m$, and a level of effort given these resources, which we denote $e_i \geq 0$. Therefore, the "performance-production function" of the individual runner is

$$\begin{aligned} &\pi_i = F(\mathbf{r}_i, e_i) \text{ with } F(\mathbf{r}_i, 0) = F(\mathbf{0}, e_i) = 0, \\ &\partial F / \partial r_{ih} \geq 0 \quad h \quad \{1, \quad , m\},\ \partial F / \partial e_i > 0 \end{aligned} \tag{1}$$ [1]

Note that, classically, we assume at this stage that resources in vector **r** exhaust all non-effort variables which determine any agent's *specific* performance: the production function F is taken to be a universal datum of human nature, somewhat analogously to what Kolm (1971, 1998) has called "fundamental preferences." F is a "fundamental transformation technology" which translates resource-effort combinations into running times.[2] We therefore assume that it is possible to explain all differences in individual "transformation capacities" as differences in certain causal variables contained in the vector **r**. What remains in the determination of the shape of F are then both universal human characteristics and exogenous data such as the weather, wind speed, etc., which affect everyone equally at given levels of effort and resources.[3]

Supposing, to begin with, that the agent is endowed with a (utility-representable) preference ordering of effort-performance bundles, we can write her utility function as

$$U_i(\pi_i, e_i) = U_i\big(F(\mathbf{r}_i, e_i), e_i\big), \text{ with } \partial U_i / \partial F > 0,\ \partial U_i / \partial e \leq 0 \tag{2}$$

Note that we assume that the runner knows with certainty what is going to be her performance given resources and effort. More important, in a model so formulated, what matters to the runner is her own *absolute* performance: what is important for her is to participate, to do the best she can with respect to herself whatever she believes the others' performances could be. This is a world of "emulation"; it is *not* the world of competition. Indeed, in a *competitive race*, what obviously matters to a runner is not her own *absolute* performance, but her *relative* performance—that is, her *rank*, which, as we will see later, might be the only instrument to obtain a *prize* (see equations (9) and (10) below).

---

[1] The zero value of the F function when either $\mathbf{r}_i$ or e_i are zero means that *both* resources *and* effort are necessary for a non-zero performance to be delivered.

[2] A purely deterministic conception of the human being would of course reject the existence of effort as pure "free will", i.e., freedom independent of any external factors. See below the discussion of Σ_i after equation (2').

[3] The exact specification of F also requires us to say something about the extent to which effort and resources are substitutes. For the sake of simplicity, we just assume here that some minimal effort is always required to be able to run. Further specifications might make sense, e.g., that particularly important differences in resources cannot be compensated for simply by differences in effort.

Accordingly, in a competitive framework, the F function does not enter explicitly as such in the utility function, but only as one determinant of the expected rank.

Call $\rho_i \in N$ the rank attained by runner i. The simplest way to model relative performance is to realise that runner i's rank depends *ex post* on each and every runner's absolute performance:

$$\rho_i = \Psi\left(\pi_1, \ldots, \pi_i, \ldots, \pi_n\right) \tag{3}$$

where n is the number of runners. However, given equation (1), runner i has to decide *ex ante* the amount of optimal effort he is going to put into the race. Writing down equation (3) as part of his decision program amounts to a very strong assumption, which physicists have come to know as "Laplace's demon": it is assumed that (*a*) there is some external viewpoint from which exhaustive causalities can be drawn up, understood, and put together to form a fully determined model of the race, and (*b*) each individual runner can take on this disengaged viewpoint *ex ante*, that is, before running. Nothing short of this—even in cases where this causal model is accepted to be stochastic—has been assumed in various versions of "perfect foresight" and "rational expectations." Equation (3), supposing we adopt it nevertheless, then allows runner i to form an expected-ranking function:

$$E^i\left(\rho_i\right) = \Psi\left(E^i\left(\pi_1\right), \ldots, \pi_i, \ldots, E^i\left(\pi_n\right)\right) \tag{4}$$

Combining this with equation (1)—which amounts to assuming that the runner knows for certain that this model of transformation applies to all his competitors as it does to himself—we obtain

$$E^i\left(\rho_i\right) = \Psi\left(F\left(E^i\left[\mathbf{r}_1\right], E^i\left[e_1\right]\right), \ldots, F\left(\mathbf{r}_i, e_i\right), \ldots, F\left(E^i\left[\mathbf{r}_n\right], E^i\left[e_n\right]\right)\right) \tag{4'}$$ [4]

We then have an expected-utility function

$$U_i\left(E^i\left(\rho_i\right), e_i\right) = U_i\left(\Psi\left(F\left(E^i\left[\mathbf{r}_1\right], E^i\left[e_1\right]\right), \ldots, F\left(\mathbf{r}_i, e_i\right), \ldots, F\left(E^i\left[\mathbf{r}_n\right], E^i\left[e_n\right]\right)\right), e_i\right) \tag{2'}$$

Each individual has her own utility function; it determines all the possible trade-offs that she can make between effort and the expected value of the race (probability of obtaining a certain rank) and thus all the possible values of her optimal effort e_i^*. However, we will realistically assume that there are some psychological constraints on

[4] It is assumed that runner i believes $\mathbf{r}_j$ and e_j to be uncorrelated.

the maximal effort the individual will be willing to perform whatever the value of the race. This means that e_i^* is bounded upward by some maximal value e_i^{MAX}. Thus, the maximisation of (2') with respect to e_i has to take into account an additional set of constraints: each runner *i* can select her effort from a *possibility set* which we will denote Σ_i. What is contained in Σ_i ? Since in the resource vector $\mathbf{r}$, "everything" has been included that is given as a datum to runner *i* and thus that she should not be held accountable for, Σ_i includes only the range of values of e_i such that $0 < e_i < e_i^{MAX}$, given the subjective (psychological) factors which affect the "preferences" between effort and its expected reward. We would like to label these factors the runner's "willingness to win."

Thus, the full optimisation program of the runner is the following:

$$\begin{aligned}&\max_{e_i \quad \Sigma_i} U_i\left(\Psi\left(F\left(E^i[\mathbf{r}_1], E^i[e_1]\right), \quad , F\left(\mathbf{r}_i, e_i\right), \quad , F\left(E^i[\mathbf{r}_n], E^i[e_n]\right)\right), e_i\right)\\ &\text{s.t. } E^i\left(e_k\right) \quad E^i\left(\Sigma_k\right) \quad k \neq i\end{aligned} \tag{5}$$

If this has a solution, it is of the form

$$e_i^*\left(\mathbf{r}_i, E^i[\mathbf{r}_1], \quad , E^i[\mathbf{r}_n], E^i[e_1], \quad , E^i[e_n]\right) \quad \Sigma_i \tag{6}$$

and it is only under the additional assumptions of (*i*) perfect information about everyone's resource vectors and (*ii*) common knowledge of everyone's utility-maximising attitude that we get a "reaction function" for *i* which can lead to a Nash-equilibrium computation of everyone's optimal effort levels:

$$\begin{aligned}&\text{Reaction function : } e_i^*\left(\mathbf{r}_1, \quad , \mathbf{r}_i, \quad , \mathbf{r}_n, e_1, \quad , e_{i-1}, e_{i+1}, \quad , e_n\right)\\ &\text{Nash equilibrium: } e_i^*\left(\mathbf{r}_1, \quad , \mathbf{r}_i, \quad , \mathbf{r}_n, e_1^*, \quad , e_{i-1}^*, e_{i+1}^*, \quad , e_n^*\right)\end{aligned} \tag{6'}$$

Equation (6') gives us the equilibrium effort level which each runner *i* would choose on the starting line given her information about the various data of the problem. From (1) this will yield an optimal running time for each *i* and hence an optimal profile of ranks, from (3):

$$\pi_i^* = F\left(\mathbf{r}_i, e_i^*\right) \tag{1'}$$

$$\rho_i^* = \Psi\left(\pi_1^*, \quad , \pi_i^*, \quad , \pi_n^*\right) \tag{3'}$$

Both these equations are simply results of the simultaneous optimising choices of all runners.

In the standard perspective which we are describing in this section, this solution can be judged *unfair* for two reasons: either all runners did not have the same possibility sets Σ_k and this kept some of them from selecting a sufficiently high level of effort ; or all runners did not have the same resource vector $\mathbf{r}_k$ and this kept their effort from being efficient enough to attain a high enough performance level. Fairness in the usual defence of competition would require at least an absence of envy, that is, no runner should prefer to have another's effort-rank combination. Let us call this criterion "U-fairness", where the "U" stands for "usual". In the terms of our notation, U-fairness requires

$$U_i\left(\rho_i^*, e_i^*\right) \geq U_i\left(\rho_j^*, e_j^*\right) \quad \forall i, j \tag{7}$$

The two conditions

$$\begin{cases} (a) \ \ \mathbf{r}_i = \mathbf{r}_j & \forall i, j \\ (b) \ \ \Sigma_i = \Sigma_j & \forall i, j \end{cases} \tag{8}$$

guarantee that (7) is satisfied, according to the classical "equal-split" argument (see, e.g., Kolm 1971): all runners have the exact same resource vector and the exact same choice set, so that the maximal available effort level is the same for everyone; therefore, any runner whose preferences are such that—given identical resources—she chooses an effort level which does not allow her to arrive at rank 1 by definition does not *want* to win given the effort this would involve. *Therefore, competition is U-fair if condition (8) is fulfilled.*

Let us look carefully at what this implies. Since by assumption the "performance technology" represented by function F is the same for all individuals, U-fairness requires essentially (a) that we equalise all internal and external resources (the vectors $\mathbf{r}$) and (b) that we provide all individuals with the same psychological opportunities for choosing their level of effort. In "concrete" terms, this means essentially two things. To satisfy condition (a) differential medical, nutritional, technical and other assistance has to be provided *ex ante* to all competitors up to the point where they have fully equalised their resource vectors. This goes as far as making all technical advances which improve resources a fully public good (see Arnsperger and De Villé, 1999 for a more detailed discussion). Moreover, along condition (b), no *differential* psychological barriers must be allowed to persist. This may again require differential *ex ante* treatment for each runner.[5]

[5] Obviously, as the whole literature on redistribution in economic environments after Dworkin (1981) has shown, deep problems can arise for the no-envy criterion because of non-transferrable inner resources (mainly, unchosen handicaps; see, e.g., Fleurbaey 1994, 1995*a*). This amounts to the maximal available effort not being the same for all runners. If (8) cannot be satisfied for physical or psychological reasons, envy-free effort-ranking profiles cannot be obtained from an equal split. Moreover, no such profiles may exist at all, i.e., solution (7) may contain no elements, even without

Notice that two things at least are, by assumption, excluded from the idea of U-fairness. First, the interdependence between preferences in determining the final ranking is not judged *per se* an element of U-unfairness. But it is true that the strategic interaction which all runners are here assumed to take explicitly into account might make such an externality irrelevant. Second, preferences themselves are assumed to be entirely under the control of the agents. Hence there is no room for any complaint on the basis of someone having "detrimental" preferences. If a runner has not won the race, it is because she has consciously decided not to do so according to her own preferences about the trade-off between effort and return which, of course, she should be held responsible for.

Provided full equalization has been carried out according to (8), competition thus appears here as a perfect screening or selection procedure which, by allowing perfect arbitrage between effort and ranking by all runners, yields a profile of ranks which *exactly reflects the differences in runners' preferences*. This, it seems to us, is the usual view underlying most arguments by those economists who want to defend competition.

2. The "eclipse" of preferences and the will to victory

We would now like to claim that this neat preference-based model of fair competition is deeply defective because it neglects two basic features of competitive settings.

First, on the starting line, even perfectly calculating runners will not carry out a trade-off between effort and ranking. They will rather become immersed in a setting where everyone acquires a *uniform maximal will to victory*—with the disturbing result that any competition geared to perfect U-fairness (that is, respecting condition (8)) will inevitably be completely non-selective: all runners must display exactly the same performance. Alternatively, if our argument below is correct, any competition which exhibits some selectivity is necessarily U-unfair. In other words, competition leads to situations in which the *real freedom* for each runner to choose her "optimal" amount of effort

---

an equal split. Alternative criteria of fairness then need to be introduced (such as "undominated diversity," for instance; see Fleurbaey 1995*b*) in order to "cover up" the impossibility of fairness—without, however, making a convincing case for these more artificial criteria. In fact, this seems to be a key defect of the whole "postwelfarist" enterprise, namely, that it tries to *recast* the notion of fairness so as to recover new *possibility* results, rather than recognising the *radical impossibility* of fairness as no-envy in all except very special situations. Indeed, it is crucial to distinguish the positive proceedings of axiomatic method, which frequently seek to uncover possibility results by adapting the content of the axioms so as to "weaken" them appropriately, and the critical proceedings of political economy (within which our present discussion is located), in which radical impossibilities may need to be tolerated as signs that something is seriously wrong with the overall socio-economic system that is being discussed.

becomes inexistent; in that sense, competition is unfair although *formal freedom*, i.e., the free will of each runner, is still present.

Second, there is another notion of freedom as fairness, which is just as important as the previous one, and which is also negated by competition. It has to do not with the freedom to choose as such, but with the fact that the constraints on the runners' choices *always* operate in the sense of *increased* competitiveness and, therefore, in the direction of an *increased* required effort on their part—the process, in fact, *never* requires constant or decreased effort on anyone's part.

2.1. *The "eclipse" of preferences*

The key to our argument lies in the *perceived structure of payments* which we claim the runners face on the starting line. In order to separate out cognitive from conceptual difficulties, let us assume that no runner has any limitations of computational ability, so that no "bounded-rationality" argument has any bite. To begin with, let us assume that the rank-to-payment mapping, which we will denote by *W*, is of the following form:

$$W : \rho \to \pi(\rho)$$
$$\begin{cases} w(\rho \mid 1 < \rho \leq n) = \varepsilon \geq 0 \\ w(1) = \Omega > \varepsilon \end{cases} \qquad (9)$$

If ε is very small, we may call this a "quasi-winner-take-all" scheme, in which the runner ranked first receives most of the prizes while all subsequent ranks are paid almost nothing. The "winner-take-all" case discussed in Frank and Cook (1995) is the particular case where $\varepsilon = 0$. For our present purposes, let us simply assume that the size of ε is such that *it prevents any of the runners from "surviving" if it is her sole source of income.* This means that, with appropriately reformulated utility functions (since now what matters to each runner is not her rank *per se* but only the *income* she gets *from* a given rank), we have

$$U_i(\varepsilon, e_i) = 0 \quad e_i \quad \Sigma_i \qquad (10)$$

Equation (10) appropriately defines "survival" as zero satisfaction, i.e., "life no longer being worth living." It deliberately gives to this notion of survival both an objective (ε) and a subjective (U_i) component. It means that each and every individual, finding herself in a situation where she gets ε, would do "anything" to get out of that situation.

Accordingly, runner *i*'s expected utility from taking part in the race is

$$\xi_i\left(e_i, E^i\left[e_1\right], \quad ,E^i\left[e_n\right]\right) = p_i\left[\rho_i = 1\right] \;\; U_i\left(\Omega, e_i\right) \qquad (11)$$
$$\partial p_i / \partial e_i > 0, \;\; \partial p_i / \partial E^i\left[e_k\right] < 0 \quad k \neq i$$

It is, however, absurd to assume that the runner will merely rest content with maximising this function by trading off between effort and probability of victory, since the alternative to winning is getting a utility of zero, i.e., simply failing to “survive.” It is well known that in such cases the only sensible assumption is *infinite risk aversion*, with the consequence that *i*’s optimisation program is simply

$$\max_{e_i} p_i\left[\rho_i = 1\right] \;\; \overset{(11)}{\quad} \;\; e_i^* = \max \Sigma_i \qquad (12)$$

Consequently, provided condition (8) has still been satisfied prior to the race, all runners will cross the finish line at exactly the same moment: since by (12) everyone will provide exactly the same (maximal) effort level, the only result of a U-fair quasi-winner-take-all race is a *total tie*. In this sense, scheme (9) brings about maximal efficiency because it forces all runners to provide maximal effort. Note that this requires no irrationality at all: each runner is fully aware that all others maximise their own probability and hence will also select the maximal effort level. In other words, quasi-winner-take-all always pushes runners towards their maximal effort, so that, if effort is upward-bounded, each of them is being driven into a corner solution. The resulting non-convexity seems to us a necessary condition for the “eclipse” of preferences.

We believe that it is a key feature of contemporary norms of economic interaction for scheme (9) to be *de facto* present in all runners’ minds; in other words, we believe that our race model can serve as a relevant metaphor for real-life problems connected with economic competition. In real-life economic competition, some firms close down, some employees lose their jobs, some salespeople do better than others: “real” competition is indeed selective. But the costs for not meeting the standards imposed by the competition are often felt to be—or actually are—unbearable for the individuals who might be subjected to them. What our analysis suggests at this stage is that these individuals’ failure to “survive” during the race should be linked not so much to inappropriate preferences for effort but to the lack of adequate compensation for unintended inequalities. In other words, actual economic competition tends to be U-unfair because, being selective but also subjected to scheme (9), it necessarily violates condition (8). *Competition submitted both to the U-fairness condition (8) and to scheme (9) is necessarily non-selective; alternatively, the only way for (12) to yield a differentiated ranking is that (8b) be violated.*

If scheme (9) is present in the form of a *will to victory* which is independent of the structure of preferences and which eventually materialises in the form of actually "non-survival" payments, the question is now: do we have reasons to believe that scheme (9) will emerge as the result of competition itself? There are at least two possible interpretations: first, a form of "psychology of shame" which focuses on a reformulation of (indirect) preferences, and second, a *sequential mechanism* by which repeated bouts of competition force individuals to gradually put aside the normal play of their preferences, given a particular but plausible setting. The two interpretations are clearly distinct: the first one locates the "winner-take-all" property in a psycho-sociological or cultural datum; the second one shows that a winner-take-all form can emerge as the end-point of a sequential process of competition. In both interpretations, however, the essential feature is the same: something is at work which *curtails individual competitors' choice sets* in such a way that each runner *perceives herself* to be facing a "winner-take-all" situation, independently of her "unconstrained" preferences. We shall first describe the two interpretations (the second one being rather longer than the first), and then focus on some of the ethical arguments which they yield concerning competition.

2.2. Preferences and the "psychology of shame"

Let us assume that social customs and norms are such that—through a process which could be justified by various socio-cultural arguments—individuals' utility functions contain a very strong "reputation effect" linked to the shame of not arriving first. This may mean, in admittedly extreme cases, that individual *i*'s utility is of the form

$$V_i\left(E^i(\rho_i), e_i\right) = U_i\left(E^i(\rho_i), e_i\right) + \Delta_i\left(E^i(\rho_i)\right),$$
$$\text{with } \Delta_i = \begin{cases} 0 \text{ if } E^i(\rho_i) = 1 \\ -\infty \text{ if } E^i(\rho_i) > 1 \end{cases} \tag{13}$$

Clearly, the "symbolic" or "reputational" component is such that psychologically the individual cannot bear to be less than first-ranked, in the sense that her expected utility becomes infinitely negative if she does not expect to arrive at rank 1. The presence of increasing psychic stress has been amply demonstrated as the result of the intensification of competitive pressures (see e.g. Lane 1991: 480-489; Aubert and de Gaulejac 1991; Schor 1992; Sennett 1998; Dejours 1998, 2000; Perilleux 2001). If individuals felt that they were able to always adjust themselves "at the margin" to the optimal effort-expected reward mix, no such stress should be observed. Thus (13) or something like it must be present if we are to account for various empirical cases of depression, burn-out, or even workplace suicide.

2.3. How competition tends to create "winner-take-all" conditions

In order to understand how a form such as (13) can come to arise in a society, we have to look at the "wider picture" and ask ourselves by what *process* the evolution of competitive conditions within a society can lead agents to *perceive* their situation as a winner-take-all one. What follows is an attempt to spell out a simple formal model in which something akin to (13) arises endogenously. What is striking is that this can be the case only if, alongside competition between runners, we introduce competition between *organisations setting up races for groups of runners*, that is, between "stadiums."

Let us suppose that there are S stadiums, denoted $s = 1, \quad, S$. At each period t (=1, 2, …) there is a race in each stadium. In a given stadium s, race t takes place given a payment scheme W_s^t characterised by a profile of nonnegative payments $\left\{w_s^t(1), w_s^t(2), \quad, w_s^t(n_s)\right\}$. There are thus n_s runners in stadium s, and we shall assume at this point that these runners do not have the possibility of moving from one stadium to another. (As we will see below, the implications of this assumption are less alarming for our present purposes than might be expected.) To simplify, let us assume that these payments always add up to a fixed amount, so that in each stadium, $\sum_{\rho=1}^{n_s} w_s^t(\rho) = \Theta_s > 0, \quad t$. The stadium finances these payments by collecting a fixed entrance fee p from each of the N_s^t spectators who come to see the race at period t. Each stadium's per-period profit is therefore

$$\Gamma_s^t\left(N_s^t\right) = pN_s^t - \Theta_s \qquad (14)$$ [6]

which depends solely on the number of spectators of the period. Spectators move from one stadium to the other in search of the lowest racing time; this means that at each

[6]An alternative specification could be $\Gamma_s^t\left(N_s^t\right) = pN_s^t - \theta N_s^t = (1-\theta)N_s^t$, where θ is the fraction of the budget paid out in "wages." This would make the distribution of wages and profits proportional to total receipts.

period, those stadiums in which the winner was not the fastest of all winners lose some spectators, while the stadium with the fastest winner gains some spectators:

$$\begin{aligned} N_s^{t+1} &< N_s^t \text{ if } \pi_s^t(1) < \pi_{\max}^t(1) \\ N_s^{t+1} &> N_s^t \text{ if } \pi_s^t(1) = \pi_{\max}^t(1) \end{aligned} \tag{15}$$

This corresponds to a particular, but to us plausible, assumption about spectator behaviour: people go to a race to see not only a winner, but a *fast* winner. Therefore, once all winners' running times are known across stadiums, a fraction of the spectators of any less-than-fastest race move to the stadium where the fastest race took place last time. An extreme case of this would be the one where *all* spectators of *all* less-than-fastest stadiums shift to the fastest one.[7] The less extreme case portrayed in equation (15) is the one where *some* of the spectators change stadiums after knowing all the winner's performances.

The effect of (15), of course, is that any stadium whose winner is not the overall fastest—let us call her the "arch-winner" of period t—is certain to lose income in the next period. So the basic idea is to motivate the stadium's own winners to perform better next time, so as to undercut the current arch-winner. Let us very crucially assume that *each stadium is run by an "owner"* whose a-priori objective is to steadily increase profits over time. This excludes the alternative scenario where the runners self-manage the stadium and organise their incentives amongst themselves—a relevant but altogether different case. What will also be presupposed is that either (*i*) there is a "shareholder value" type of constraint which forces net profits to increase over time, or (*ii*) the owners figure that it is rational for them to try to increase profits each period, since they foresee they are unlikely to have the arch-winner every period. Therefore, we shall assume that the stadium owners' objective for any period t is

$$\tilde{\Gamma}_s^t - \tilde{\Gamma}_s^{t-1} > 0 \tag{16}$$

$$\Leftrightarrow \; \tilde{N}_s^t > \tilde{N}_s^{t-1} \text{ given } p \text{ and } \Theta_s \tag{16'}$$

[7] In that case, in the next period S–1 stadiums have to organise a race with no spectators at all, and no receipts, in the hope that their winner this time will be the fastest, so that all spectators next time will flow to them, and so on.

This is, of course, an objective in terms of *notional* or a-priori profit, which we indicate by the tildas (as opposed to realised quantities, which are without a tilda); it means that, notionally (i.e., before having observed the effective running times of all winners in all stadiums) the owners want to have an increase in spectatorship and they want to motivate their runners in that direction. Since running times are perfectly observable, there is no problem of preference revelation, hence no principal-agent problem. Given the arbitrage rule (15) and the runners' motivational structure (10)-(11), and given distributive assumption (8), the owner's objective is to try to elicit increased performance (in the hope of having the arch-winner in one's own stadium) by announcing an increased reward in case of victory. Thus, the implication of (16) is that the distribution of rewards has to become less and less egalitarian through time, so as to sustain profit growth by eliciting ever higher performance on the part of the winner. In the absence of knowledge concerning individual preferences, and given condition (8), this means that *all* runners' performances have to be boosted over time. There at least two (and perhaps more) ways of trying to elicit such steadily increased performance in the hope of "screening out" the arch-winner at least some of the time. One way is to offer a *premium* to the winner in case she undercuts last period's arch-winner, which implies that all non-winners get less if this happens. We may assume (although this is not strictly necessary) that the rule does not apply to the stadium in which last period's arch-winner was located:

$$
\begin{aligned}
& s : \pi_s^t(1) < \pi_{\max}^t(1), \\
& \begin{cases} w_s^{t+1}(1) > w_s^t(1) \text{ if } \quad \pi_s^{t+1}(1) > \pi_{\max}^t(1) \\ w_s^{t+1}(1) = w_s^t(1) \text{ if not} \end{cases}
\end{aligned}
\tag{17}
$$

Notice that, by assumption, this implies that all $w_s(\rho > 1)$ decrease over time if there is undercutting. We may assume that the aggregate decrease in the non-winners' rewards is distributed equally over all non-winners in each period. Equation (17) poses a classic problem of potential ineffectiveness of the mechanism: if all runners together agree to not running faster than last period's arch-winner, no premium will be paid because the mechanism is conditional on results. Of course, such a "cartel" is highly unstable in the absence of its formalisation as a "union," as any runner has an incentive to deviate and try to outperform last period's arch-winner, in which case that one runner gains but all the $n_s - 1$ others stand to lose. If all runners fear this without any possibility for individually binding commitment, there will be generalised effort towards outperforming the arch-winner. Moreover, even if there is a "cartel" and no one deviates, this may in the long run be detrimental to all runners because unless the arbitrage rule (15) is altered, the stadium will eventually lose all its spectators. If all runners internalise this longer-run perspective (which is not the case in the preferences as given by (10)-(11)), then (17) will be an effective mechanism.

The other method to try and elicit undercutting is simply to announce unconditional increases in reward inequality:

$$\begin{cases} w_s^{t+1}(1) > w_s^t(1) & s : \pi_s^t < \pi_{\max}^t \\ w_s^{t+1}(1) = w_s^t(1) & s : \pi_s^t = \pi_{\max}^t \end{cases} \tag{18}$$

This differs from (17) in only one respect, but one which is crucial: the conditionality is not on willingness to improve over the past arch-winner; rule (18) says only that in all stadiums where the winner is not also the arch-winner, the next winner's reward is automatically increased—at the expense of all other runners' rewards.[8] We assume this is not the case for the arch-winner herself, but again this is not absolutely essential. Mechanism (18) offers no way of creating a "cartel" because rewards are pulled apart independently of any increase in the winners' performances.

The reason we discuss these mechanisms is to convey a simple fact. Although the precise dynamic path of performances inside and across stadiums is immensely complex and depends on all individual runners' preferences for income and effort (equation (10)) as well as on their appraisals of probabilities of winning any given race, what is certain is that overall the system is going to witness an eventual increase in performances (since we assume that all preferences are monotonic in income), coupled with a steady widening of the gap between top and less-than-first-ranked rewards. What is going to happen in case the stadium owners all apply rule (18)? The steady state will be attained when in all stadiums the rewards $w_s(\rho > 1)$ have fallen below a level $\varepsilon > 0$ such that no one is able to materially survive—whereas in each stadium the reward for winning next period's race has grown as large as allowed by the stadium's constant-budget limit: $w_s^{t+1}(1) = \Theta_s - (n_s - 1)\varepsilon$.

Thus, in steady state, the interaction between the spectators' arbitrage behaviour and the stadium owners' incentive mechanism yields, for each stadium, a situation exactly identical with (9) above. In this case, fierce competition *between* stadiums leads to the

[8] This device is closely related to the now well-studied incentive mechanisms of the "economics of tournaments." The seminal reference is Lazear and Rosen (1981). For the interested reader, a general survey of this topic is provided in *The New Palgrave*, volume 2, pp. 745-746.

point where competition *within* each stadium degenerates into a purely non-selective process: as we saw earlier, faced with the prospect of non-survival in case she doesn't arrive first, each runner will apply (11) and simply "run for her life." In the limit, faced with this winner-take-all situation, no runner any longer has any possibility of making a trade-off between effort and income: given the fact that condition (8) is satisfied by assumption, the risk—created by competition between stadiums and the need to keep profits growing—of having no chance of survival at all if one does not match the arch-winner *eventually transforms each and every runner into a local winner, and eventually into an arch-winner.* In other words, the limit of our process of competition between stadiums makes it *look like* all runners have acquired *a uniform "will to victory"*: they run as fast as they can and, having the same initial conditions, they arrive at the exact same moment. Of course, we assume that the *per capita* reward for winning, which by definition is Θ_s/n_s, is greater than the survival level ε of each runner.

The end-state of the competitive process is therefore a state where, in every still existing stadium (some may have been eliminated through too long spells without an arch-winner), there are no longer any winners or losers—this is the precise form which our paradox takes here, namely that when *U-fair* and *efficient* (in the sense of inducing maximal effort on the part of each and everyone runner), the competition within each stadium is bound to exhaust itself as a process whereby the fittest are to survive, i.e., to be selected.

3. Ethical implications: The curse of single-minded effort and the criterion of S-fairness

"So what?" might be the reaction of many economists. "Haven't you constructed a model in which, through an arbitrage mechanism, competition eventually selects the most efficient, i.e., effort-performing, agents? What does it matter if all runners in the *still existing* stadiums all arrive at the same time? Haven't we selected the fastest stadiums nevertheless? In fact, is there any issue of fairness left to discuss, given that all runners have exactly the same chances at the start of each race *and* that they all arrive at exactly the same moment?"

Our reply to this is that such a diagnosis overlooks a much deeper, and virtually invisible source of unfairness. If we look more closely, the very process by which the non-selectivity of U-fair competition is reached contains a stark element of *sequentially emerging unfairness*, which we shall call *S-unfairness* (where the "S" stands for "sequence"): pressured by the spectators' demands for the fastest possible race, all stadium owners send their runners into a spiral whereby their latitude for freely expressing their trade-off between income and effort within acceptable bounds steadily decreases. True enough, it is very difficult to say at which point in the sequence this latitude *becomes* unfairly restricted; what is certain, however, is that in the *steady state* such latitude is totally absent: each runner applies rule (12) because of the absolute impossibility of doing anything else. Although U-fair, in the distributive sense of condition (8) being satisfied by assumption, competition becomes S-unfair because it gradually restricts freedom *due to the spectator-driven sequence whereby a flow of increasing profit over time means eliciting ever more single-minded—or, for that matter, absent-minded—effort from each and every runner*. This being pushed into single- or absent-minded effort, regardless of one's preferences, is an element of the model which should not be overlooked: it points towards a form of *alienation* which we believe must not be kept out of the fairness-of-competition debate.

While U-fairness has to do mainly with a distributive norm to be satisfied at the outset of each bout of competition (of each "race"), S-fairness has to do with each competitor having the real freedom to trade off effort against leisure. It is not, in itself, an instantaneous distributive criterion: it has to do with the sequential evolution of choice sets, which we claim can be ethically objectionable even if—as is the case in our paper—U-fairness is assumed to hold at all stages in the process. Of course, if—as is the case in the real world—U-fairness itself fails to be satisfied, the unfairness of competition is compounded; but this does not imply that, analytically and ethically speaking, the two kinds of unfairness should not be carefully kept apart. They "communicate" through the fact that, as we have shown, the very emergence of S-unfairness makes U-fairness a trivial, and hence ethically empty, criterion. To put it differently, U-fairness, which is the pet argument of most defenders of competition, is a relevant criterion only as long as competition is S-fair; but since those who focus on U-fairness usually completely neglect S-unfairness, their very focus on their pet argument makes it less and less relevant for what they want to argue, namely that competition is supposedly an efficient screening device.

Of course, one immediate criticism is that we have assumed runners to be immobile across stadiums. Could not many of them find a way out of the "Satanic mill" by changing stadiums, for instance by moving to a stadium where the structure of rewards is not yet too uneven? The reply is relatively straightforward: of course they can, but this mobility will only hasten the arrival of the steady state because stadium owners will only take in the *faster* runners in the hope of taking in a potential arch-winner. As a result, the most mobile runners will be those who already had a relative preference for effort in their past stadium: this will compel all other stadiums to increase the differentials in their reward structures even more so as to counteract the receiving stadium's new recruitment. As a result, while the *dynamics* of the system will surely be quite strongly affected by mobility across stadiums, the end-point is unchanged: some runners will be able to provisionally shelter themselves against too much restriction on their tradeoffs, but in the steady state all possibilities of mobility will have been exhausted. The reward for winning is the same everywhere, as is the gap compared to non-winning rewards.

So consumer arbitrage and the need to keep up profits sends the system towards a limit-situation that is akin to the Walrasian equilibrium in neo-classical economics (see e.g. Debreu, 1959 or Arrow and Hahn 1971): opportunities for heightening performance, and hence for keeping profits on an increasing path, have been exhausted up to the point where there is literally no competition going on anymore. In the present model, competition has *exhausted itself* at the cost of a *hidden unfreedom* on the part of each runner and of a *lid on profitability* for each individual stadium—and rather in the same way, the expression "competitive equilibrium" is a misnomer: it is precisely the situation where no competition is no longer happening because pure profits are down to zero and no consumer's welfare can be increased without decreasing anyone else's. Therefore, "general competitive analysis" is ironically the analysis of a world in which every participant in the competition (every "producer") performs exactly the same, and in which there is no competition (see De Villé 1980). This may be felt to be merely a recast of old "Austrian" criticisms against the notion of *equilibrium* in economics, and indeed it is to some extent—but what the Austrians such as von Mises, Hayek or Kirzner never admitted was the alienation of those who participate in the *process* of (out-of-equilibrium) competition, and this particular source of unfairness needs to be brought back into the debate.

References

Arnsperger, Ch., and Ph. De Villé (1999), “La concurrence comme défi à l’éthique des affaires: Quand les relations sociales sont en jeu”, *Éthique Publique* **1** (2), October 1999, 32-41.

Arrow, K. J., and F. Hahn (1971), *General Competitive Analysis*, San Francisco: Holden Day.

Aubert, N., and V. de Gaulejac (1991), *Le Coût de l’excellence*, Paris: Éditions du Seuil.

De Villé, Ph. (1980), “Time and Dynamics in Economics: The Lessons from General Equilibrium Theory”, in B. H. Banathy (ed.), *Systems Science and Science*, Society for General Systems Research, Proceedings of the 24th North-American Meeting, Louisville: Systsmes Science Institute, 260-265.

Debreu, G. (1959), *Theory of Value. An Axiomatic Analysis of Economic Equilibrium*, New Haven: Yale University Press.

Dejours, Ch. (1998), *Souffrance en France: La banalisation de l’injustice sociale*, Paris: Seuil.

Dejours, Ch. (2000), *Travail, usure mentale: Essai de psychopathologie du travail*, second edition, Paris: Bayard.

Dworkin, R. (1981), “What is Equality? Part 2: Equality of Resources”, *Philosophy and Public Affairs* **10**, 283-345.

Fleurbaey, M. (1994), “L’absence d’envie dans une problématique post-welfariste”, *Recherches Economiques de Louvain* **60**, 9-41.

Fleurbaey, M. (1995*a*), “Three Solutions to the Compensation Problem”, *Journal of Economic Theory* **65**, 505-521.

Fleurbaey, M. (1995*b*), “Equal Opportunity or Equal Social Outcome?”, *Economics and Philosophy* **11**, 25-55.

Frank, R. H., and Ph. J. Cook (1995), *The Winner-Take-All Society*, New York: Free Press.

Kolm, S.-Ch. (1971), *Justice et équité*, Paris: Éditions du CNRS.

Kolm, S.-Ch. (1998), “Preface 1998” to the English translation of Kolm (1971), Cambridge, Mass.: MIT Press.

Lane, R. (1991), *The Market Experience*, Cambridge: Cambridge University Press.

Lazear, E. P., and S. Rosen (1981), “Rank-Order Tournaments as Optimum Labor Contracts”, *Journal of Political Economy* **89**, 841-864.

Perilleux, T. (2001), *Les Tensions de la flexibilité*, Paris: Desclée de Brouwer.

Schor, J. (1991), *The Overworked American: The Unexpected Decline of Leisure,* New York: HarperCollins.

Sennett, R. (1998), *The Corrosion of Character: The Personal Consequences of Work in the New Capitalism*, New York: Norton.