Robust rankings of multi-dimensional performances
An application to Tour de France racing cyclists*

Laurens Cherchye† and Frederic Vermeulen‡

September 24, 2004

Abstract

There is a general interest in ranking performances (*e.g.*, in sports or policy), which essentially implies aggregating several performance dimensions. The usual approach considers a 'cardinal' linear weighting of the different single-dimensional performance indicators. We present an alternative approach, which merely requires 'ordinal' information regarding the importance of the different performance dimensions. We argue that this approach is robust with respect to alternative specifications of the (possibly non-linear) underlying performance aggregation function. An application to Tour de France racing cyclists (in the period 1953-2004) illustrates the approach. We find that Eddy Merckx, Bernard Hinault and Lance Armstrong (robustly) dominate almost all other racing cyclists in our sample, while they do not dominate each other. A net-dominance metric ranks Bernard Hinault on the first place in our sample; Eddy Merckx and Lance Armstrong follow very closely *ex-aequo* on the second place.

*We thank Carl Berteele, Renaat Schotte, Catherine Van Eylen and Karl Vannieuwkerke for their willing cooperation.

†University of Leuven, Campus Kortrijk and Center for Economic Studies; Fund for Scientific Research - Flanders (FWO-Vlaanderen). E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kulak.ac.be.

‡Tilburg University, CentER. P.O. Box 90153, NL-5000 LE Tilburg, The Netherlands. E-mail: frederic.vermeulen@uvt.nl.

1 Introduction

There is a general interest in performance rankings. This is most apparent in a sports context; see, *e.g.*, the widespread use of football team rankings, baseball team rankings, and so on. But the practice of performance ranking actually covers a much wider field. For example, the Joint Research Centre (2004) of the European Commission reports on cross-country policy performance rankings in the fields of environment, sustainability, economy, technological development, *etc*. Another example is provided by the Human Development Index, which measures a country's performance in basic dimensions of human development, like life expectancy, education level and standard of living (Fukada-Parr *et al.*, 2004). Finally, there are the well-known rankings of economists and economics departments, based on diverse measures of journal output (*e.g.*, Coupé, 2003; Kalaitzidakis *et al.*, 2003; Lubrano *et al.*, 2003).

Essentially, the production of performance rankings boils down to aggregating several performance dimensions. The usual approach uses an aggregation statistic that can be represented as a weighted sum of the indicators associated with each single performance dimension. In general, using c_{ij} to denote observation i's ($i = 1, ..., n$; with n the total number of observations) performance in terms of the j'th ($j = 1, ..., k$; with k the total number of performance dimensions) positive-valued indicator, the aggregate index for each i can be written as:

$$I_i^c = \sum_{j=1}^{k} w_j c_{ij}, \tag{1}$$

where w_j is the weight associated with each single-dimensional indicator j. For simplicity, we further assume that each performance dimension can be considered as a 'good', such that the associated weights w_j are always positive.

The interpretation of the index in (1) is obvious: better performance in a particular dimension generally implies a higher overall index value. Moreover, the higher the weight of that dimension, the higher the increase of the index value. An advantage of such a weighted sum index is that it provides a complete ranking: it allows for comparing any two observations $i_1, i_2 \in \{1, ..., n\}$ in terms of their overall/aggregate performance. It is also clear, however, that different weighting schemes (even if they preserve the same ordering) will yield other index values; and, importantly, more often than not the most appropriate weighting scheme is not readily available. In turn, this makes that the concomitant (complete) ranking will not be generally robust. In addition, the linear aggregation function implicitly assumes a constant trade-off between the different performance dimensions, which may often be problematic. For example, it is frequently the case that for an observation $i \in \{1, ..., n\}$ the relative value/weight of dimension j_1 as compared to j_2 ($j_1, j_2 \in \{1, ..., k\}$) increases when c_{ij_1} decreases relative to c_{ij_2} (see, *e.g.*, the common assumption of diminishing marginal rates of substitution in economics). More generally, rankings on the basis of linear aggregation functions are not robust with respect to often more realistic non-linear representations of the implicit performance production trade-offs.

We propose a methodology for obtaining performance rankings that are robust both in terms of the weighting scheme that is employed and in terms of the (possible non-linear) trade-offs between the different performance dimensions. Its philosophy is closely related to what is commonly referred to as Data Envelopment Analysis (DEA; see Cooper *et al.*, 2000, for a methodological introduction). As argued, *e.g.,* by Melyn and Moesen (1991), DEA applies a 'benefit-of-the-doubt' weighting, which essentially means that an observation is outperformed by other observations only if there does not exist a weighting scheme such that it obtains the best overall performance. This obtains performance rankings that are robust with respect to the specific weighting scheme that is used: an observation is DEA-dominated if *for each possible weighting scheme* (*i.e.*, including the 'most favorable' scheme) there is another observation that obtains a higher aggregate index value. DEA has recently been applied within a sports context (*e.g.*, Lozano *et al.*, 2002; Haas *et al.*, 2004; De Oliveira and Callum, 2004; Einolf, 2004) and within a cross-country policy evaluation context (*e.g.*, Cherchye, 2001; Evans *et al.*, 2003; Cherchye and Kuosmanen, 2004; Cherchye *et al.*, 2004).

Even though it has an attractive weighting interpretation, the DEA approach still applies a linear (*i.e.*, weighted sum) aggregation of the different performance dimensions. Our methodology drops this linearity (or constant trade-off) assumption. It merely uses information regarding the ordering of the different performance dimensions. Indeed, while it may often be difficult to specify *how much* one performance dimension is more important than another dimension, it is usually fairly easy to determine simply *that* the first dimension is more important than the second. Putting it differently, 'ordinal' weighting information is often much easier to obtain than 'cardinal' weighting information. Although such an ordinal ranking may also be somewhat arbitrary, it is clearly more robust than associating cardinal weights to the criteria; *e.g.*, a given ordinal ranking encompasses all possible cardinal rankings that are order preserving. Our methodology then obtains robust performance rankings from such ordinal information by implementing an intuitive 'compensation principle'. Specifically, the method identifies for each observation (i) the set of other observations that it outperforms and (ii) the set of observations that outperform it; these sets apply to any (possibly non-linear) aggregation function that respects the aforementioned compensation principle. Conveniently, given that the method is based on the intuitive compensation principle, it is easy-to-communicate. In addition, as we will show, the method is easy to implement computationally.

An application to Tour de France racing cyclists illustrates our methodology for ranking sports performances. Indeed, a yearly recurring question, mostly somewhere in July, is which racing cyclist can be considered as the best Tour de France participant ever. Answering this question is somewhat tricky; the overall evaluation of the Tour de France performance is not obvious at all. For example, if we solely focus on the number of final victories, then Lance Armstrong would outperform all others (including the five times winners Jacques Anquetil, Eddy Merckx, Bernard Hinault and Miguel Indurain) because of his sixth Yellow jersey (which indicates the leader in the general time classification after the final stage in the Tour) in 2004. As soon as other dimensions are taken into account, however, the question becomes substantially more difficult. For example, how do we deal with the fact that Merckx and Hinault each won the Green jersey (which indicates the leader in the points classification after the final stage) at least once, whereas Armstrong never managed to do so? And what about Merckx's 34 individual day victories as compared to Armstrong's 21? Generally, answering the question depends on which criteria are taken into account and on the weight that is accorded to each of these criteria. Our robust approach turns out to be a useful tool for addressing the issue.

The remainder of the paper unfolds as follows. The next section introduces the methodology, and argues its robust nature. Section 3 presents the data and empirical results of our application to Tour de France racing cyclists. Section 4 concludes.

2 Methodology

For simplicity, we assume in the following that the different performance indicators are ranked in *descending* order according to their importance. That is, dimension 1 is at least as important as dimension 2, which is in turn at least as important as dimension 3, and so on. As argued in the introduction, this is the sole (ordinal) information that is used in our robust ranking procedure. Our methodology builds on a compensation principle that includes this information in the performance ranking. Using V to represent the performance aggregation function that implicitly underlies the ordering (but that is not observed), we may define this principle as follows:[1]

Definition 1 *An aggregation function* $V : \Re^k \to \Re$ *satisfies the* compensation principle *if for* $i_1, i_2 \in \{1, ..., n\} : \sum_{j=1}^{\overline{k}} c_{i_1 j} \geq \sum_{j=1}^{\overline{k}} c_{i_2 j} \; \forall \overline{k} \in \{1, .., k\} \Rightarrow V(c_{i_1 1}, ..., c_{i_1 k}) \geq V(c_{i_2 1}, ..., c_{i_2 k})$.

In words, the compensation principle states that one unit more of a higher ranked performance dimension may compensate for one unit less of a lower ranked dimension, but not *vice versa*. Consequently, an observation $i_1 \in \{1, ..., n\}$ can only dominate another observation $i_2 \in \{1, ..., n\}$ if i_1 performs at least as good as i_2 in terms of the highest ranked performance dimension (*i.e.*, $c_{i_1 j} \geq c_{i_2 1}$). Next, when regarding the second dimension, even if i_1 does not perform as good as i_2 in terms of this dimension, dominance of i_1 can be obtained if a better performance in the more important dimension 1 compensates this worse performance in dimension 2. Formally, this means that the sum of i_1's performance indicators 1 and 2 should not be below the same sum for i_2 (*i.e.*, $c_{i_1 1} + c_{i_1 2} \geq c_{i_2 1} + c_{i_2 2}$). And so on. Intuitively, the compensation principle directly exploits the (sole) ordinal information that is available. It is worth noting that this intuition effectively applies only if the different performance indicators are expressed in a comparable measurement unit. This is indeed the case in our application (as well as in many other settings where similar ranking issues occur).

The compensation principle immediately institutes the following dominance metric (for any pair of observations $i_1, i_2 \in \{1, ..., n\}$):

$$I^o_{i_1, i_2} = \min_{\overline{k} \in \{1, ..., k\}} \left(\sum_{j=1}^{\overline{k}} (c_{i_1 j} - c_{i_2 j}) \right). \tag{2}$$

We say that observation i_1 dominates i_2 if $I^o_{i_1, i_2} \geq 0$; it means that $V(c_{i_1 1}, ..., c_{i_1 k}) \geq V(c_{i_2 1}, ..., c_{i_2 k})$ for any possible performance aggregation function that satisfies the compensation principle. For some given specification of the importance ordering, the metric is easily computed from the available data on the single-dimensional performance indicators. By simply using the intuitive and generally acceptable compensation principle, we obtain pairwise dominance relationships that most directly let 'the data speak for themselves'.

[1] We note that this compensation principle is similar in spirit to the 'sequential dominance' idea in the literature on welfare comparisons (*e.g.*, Atkinson and Bourguignon, 1987).

Using the metric in (2), we can illustrate the robustness of our dominance-based ranking procedure with respect to alternative specifications of the aggregation function V. Specifically, it includes the following two limiting scenarios: (i) all weight is accorded to only a single performance dimension; and (ii) all performance dimensions get the same weight. First, we have that $c_{i_1 1} < c_{i_2 1}$ implies $I^o_{i_1,i_2} < 0$. In other words, we can never specify that i_1 dominates i_2 if the former observation is dominated by the latter in the single most important performance dimension; this makes that the dominance relationships obtained from the metric in (2) are always consistent with those associated with the extreme scenario (i). Next, $\sum_{j=1}^{k} c_{i_1 j} < \sum_{j=1}^{k} c_{i_2 j}$ equally entails $I^o_{i_1,i_2} < 0$; this immediately obtains consistency with the dominance relationships corresponding to the extreme scenario (ii). Clearly, the same robustness argument extends to all intermediary (possibly non-linear) specifications of V.

To illustrate our procedure, we consider the fictitious example in Table 1, which includes data on the number of Yellow jerseys and Green jerseys for four (imaginary) Tour de France racing cyclists. Suppose now that we attach a greater weight to the Yellow jersey than to the Green jersey (while we do not specify the exact trade-off value). Clearly, racing cyclist A dominates both racing cyclists B and C: he performs at least as good with respect to both criteria. However, although D gained the Green jersey four times, compared to two times for A, the latter racing cyclist also dominates the former, since his three Yellow jerseys compensate for three Green jerseys of racing cyclist D. In the same way, it is easily checked that also B dominates C. However, neither of these latter two racing cyclists dominates racing cyclist D; B and C do not have enough Yellow jerseys to compensate for the additional Green jerseys of D. Table 2 summarizes who (see column title) is dominating who (see row title) in this example.

So far, we have focused on pairwise dominance relationships, which imply a ranking for any combination of two observations. In some instances, it may be interesting to construct a full ranking of all observations in the sample, starting from this pairwise dominance information. We see different possibilities. One natural candidate is the difference between (i) the number of other observations that a given observation dominates and (ii) the number of other observations that dominate this evaluated observation; higher values of this 'net-dominance' metric then correspond to a higher ranking within the full sample. In our fictitious example, this procedure would rank cyclist A on the first place, followed by -maybe somewhat unexpectedly- the cyclists B and D and, finally, cyclist C. Evidently, other ranking procedures are equally possible. For example, one may focus exclusively on the number of observations that is effectively dominated or, conversely, the number of observations that is dominating each individual observation. While our following application will mainly concentrate on the net-dominance metric (for the sake of compactness), we stress that it may be insightful to simultaneously consider alternative ranking procedures in practice; each ranking will yield information that is interesting on its own.

[Table 1 about here]

[Table 2 about here]

3 Application to Tour de France racing cyclists

We apply the just described methodology to a sample of modern Tour de France participants, hereby focusing on the following six single performance indicators: the number of Yellow jerseys (allocated to the leader in the general time classification after the final stage in the Tour), the number of second places in the general time classification (after the final stage), the number of third places in the general time classification (after the final stage), the number of Green jerseys (leader in the general points classification after the final stage), the number of Red Polka Dot jerseys (leader in the best climber classification after the final stage) and the number of individual day victories.[2] In order to determine the ordinal ranking that is needed for the exercise, we consulted sports journalists of the Flemish public broadcasting service. This eventually obtained the following ranking: (1) Yellow jersey, (2) second place, (3) third place, (4) day victory, (5) Green jersey and (6) Red Polka Dot jersey.[3] The ranking criterion that we focus on is the net-dominance metric introduced in the previous section (*i.e.*, the difference between the number of dominated cyclists and the number of cyclists by whom one is dominated).

Our sample includes all racing cyclists who were able to 'score' on one of the above six dimensions in at least one of their participations to the Tour

[2] Note that after each stage in the Tour de France, each of the three jerseys are handed out to the respective leaders in the different classifications. We only focus on the final winners of the classifications (*e.g.*, the number of Yellow jerseys is here the number of times a racing cyclist was able to win the general time classification after the last stage). Next, day victories in team trials are not taken into account.

[3] Different journalists gave slightly different ordinal rankings. We have chosen the ranking that most closely resembles the journalists' answers. Of course, the exercise can easily be replicated for alternative rankings. Interested readers may contact us to get results for their personal ranking.

de France. Since the Green jersey was only introduced in 1953, we restrict attention to racing cyclists participating to the Tour between 1953 and 2004 (both included).[4] As a consequence, our following analysis excludes excellent racing cyclists like Philippe Thys (who obtained three final victories) or Fausto Coppi and Gino Bartali (who are generally considered as belonging to the circle of best racing cyclists ever). However, it should be stressed that all racing cyclists who won the Tour five or more times are included in the analysis.

The full sample contains 499 racing cyclists; all data were obtained from L'Equipe (2002) and Société du Tour de France (2004). Table 3 presents summarizing frequency information for the six indicators. We find that only one racing cyclist was able to win the Tour de France six times (namely Lance Armstrong); and four other racing cyclists managed to win the Yellow jersey five times (namely Jacques Anquetil, Eddy Merckx, Bernard Hinault and Miguel Indurain). We obtain closely similar frequency distributions for the second and third places, and the Green and Red Polka Dot jerseys. Record holders with respect to these criteria are respectively Joop Zoetemelk (six second places), Raymond Poulidor (five third places), Erik Zabel (six Green jerseys) and Richard Virenque (seven Red Polka Dot jerseys). Evidently, the frequency table for the individual day victories looks somewhat different. Most of the racing cyclists in the dataset at least won one stage in the Tour (only 17 individuals have no single day victory). The record holder is Eddy Merckx with 34 victories, followed by Bernard Hinault (28 victories) and André Darrigade (22 victories). As these simple key figures suggest, racing cyclists like Merckx and Hinault perform well on several criteria. However, no participant outperforms the other racing cyclists in every single dimension, which makes aggregating the different performance dimensions an interesting question.

[Table 3 about here]

Table 4 presents a selection of dominance relationships for our dataset. For compactness, we only tabulate a fraction of the full (499×499) dominance matrix.[5] Specifically, the columns of Table 4 correspond to all the racing cyclists who won the Tour de France at least three times. The rows then

[4] In 1975, the White jersey was introduced for the best young rider in the Tour de France. It replaced the award for the winner of the so-called Combination classification (combining three or four classifications in the Tour de France). Since the rules with respect to these criteria frequently changed in the history of the Tour de France, we only focus on the aforementioned six criteria.

[5] The full dominance matrix can be obtained from the authors at simple request. This also applies to the results in Table 5.

contain the cyclists who won the Yellow jersey at least one time, and the record holders for the other performance dimensions (we note that some Yellow jersey winners also obtain the highest values for second places and day victories).

Just like in the fictitious example above, entries in the table show the (row) racing cyclists that a particular (column) racing cyclist dominates. To illustrate the interpretation of this dominance table, we discuss a number of interesting patterns that emerge. First, when looking at the racing cyclists who won the Tour de France at least five times, we find that Armstrong, Hinault and Merckx are only dominated by themselves. Even though Armstrong is the record holder for the number of Yellow jerseys, the fact that he has one more final victory than Hinault and Merckx cannot compensate for the other criteria on which the latter two cyclists scored better (notably: second places, day victories and Green and Red Polka Dot jerseys; see Table 5). On the other hand, we observe that Anquetil is dominated by all four racing cyclists who were just cited (including himself). Although Anquetil obtained a strictly higher number of third places than Armstrong, Hinault and Merckx (see Table 5), the latter three dominate the former at the overall level since they performed strictly better on other, more important criteria; their performance in terms of these more important dimensions thus compensates for their worse performance in terms of the total number of third places.

Next, it is worth considering the last three rows in the table, which contain the record holders for the number of third places (Raymond Poulidor), Green (Erik Zabel) and Red Polka Dot jerseys (Richard Virenque). Quite strikingly, no racing cyclist (except from himself) dominates Poulidor. The reason is that Poulidor obtained no less than three second places and five third places. No other cyclist can compensate this by higher scores in terms of the other dimensions that are at least as important. For example, although Armstrong's higher ranked six final victories compensate for Poulidor's three second places and for three of the five third places, they cannot compensate for Poulidor's remaining third places. Since the number of third places is more important than the number of day victories, it is not possible to compare Armstrong's 21 victories to Poulidor's performance that is not yet compensated for (notably two remaining third places and seven individual day victories). As for this specific case, it is important to stress that this does not mean that Poulidor himself dominates cyclists like Armstrong, Merckx and Hinault; none of the items where Poulidor outperforms these three can compensate their dominance in terms of the most important criterion.[6] Fi-

[6] More generally, the fact that 'observation a does not dominate observation b' should

nally, the table shows that Richard Virenque and Erik Zabel are dominated by most of the column racing cyclists, but not by all of them. Virenque, for example, is not dominated by Greg LeMond and Louison Bobet, who each won three Yellow jerseys. Erik Zabel (six Green jerseys and 12 day victories) is even not dominated by Miguel Indurain, who won the Tour five times.

[Table 4 about here]

Table 5 provides a ranking of modern Tour de France racing cyclists based on the net-dominance metric (computed with respect to the full sample of 499 cyclists). For completeness, we also report the number of cyclists that each evaluated racing cyclist dominates (see the column 'Dominating') and the number of racing cyclists by whom the evaluated individual is dominated (see the column 'Dominated'). To save space, we restrict to tabulating the first 30 racing cyclists in the full ranking. The net-dominance metric indicates Bernard Hinault as the best modern Tour de France participant: he dominates 496 (of the 499) racing cyclists, while he is only dominated by himself. Hinault is closely followed by Lance Armstrong and Eddy Merckx, who end *ex-aequo* on the second place of the ranking: they both dominate 495 participants and are only dominated by themselves. (Remark that an equal ranking here does not mean that Armstrong and Merckx dominate each other.) The results also indicate that one does not need many final Tour victories in order to be highly ranked. The case of Joop Zoetemelk, who won the Yellow jersey only once, is exemplary in this respect: thanks to six second places and ten day victories, he dominates 485 racing cyclists, while he is dominated by only two participants (*i.e.*, Bernard Hinault and himself). Erik Zabel closes the list of the 30 most highly ranked Tour de France racing cyclists: although he never stood on the main podium in Paris, his six Green jerseys and twelve day victories make him effectively dominate 421 participants, which is much above the number of seven racing cyclists who dominate him.

[Table 5 about here]

---

not imply that 'observation b dominates observation a'. In this respect, it is worth recalling that we present a procedure for performance ranking that is robust with respect to the (cardinal) specification of the performance aggregation function. Specifically, we only retain dominance relationships that hold *for all possible aggregation functions* that obey the compensation principle in Definition 1.

4 Conclusion

We have presented a methodology for robust ranking of multi-dimensional performances. It builds on an intuitive compensating principle, which essentially states that better performance in more important performance dimensions can compensate for worse performance in less important dimensions, but not *vice versa*. Using this, the method only needs ordinal information regarding the importance of the different performance dimensions, so avoiding the often difficult and controversial (cardinal) specification of a performance aggregation function. As such, it lets 'the data speak for themselves' in a most genuine sense. This obtains pairwise dominance relationships that are robust to any possible specification of the aggregation function that preserves the specified dimensional ordering. Interestingly, the underlying compensation principle is easy-to-communicate and the dominance relationships are easy-to-compute, which makes the presented approach attractive from a practical point of view.

The pairwise dominance information can be used for obtaining a full ranking of the different observations in the sample under study; for example, we have suggested a net-dominance metric as a useful tool for such a purpose. Still, it is worth to stress at this point that the pairwise relationships as such may already provide useful information; *e.g.*, in an interactive setting, dominating observations may constitute useful benchmarks for dominated performance units that pursue to improve their own performance. Finally, while we have applied the method to ranking sports performances, we believe that it may also be a valuable instrument in other contexts (*e.g.*, policy benchmarking in a cross-country policy assessment or an evaluation of economics departments; see the applications referred to in the introduction).

Our application to Tour de France racing cyclists demonstrates the practical usefulness of this robust ranking approach. For example, it shows how the pairwise dominance relationships can be useful for discriminating between the five times Yellow jersey winners (*e.g.*, Eddy Merckx and Bernard Hinault versus Jacques Anquetil and Miguel Indurain). Also, it enables robust positioning of the six time winner Lance Armstrong *vis-à-vis* these different five times champions; and it allows for comparing multi-winners in the different dimensions that we consider (Yellow Jerseys, second places, third places, day victories, Green jerseys and Red Polka Dot jerseys). Finally, our net-dominance metric obtains a full ranking of the modern Tour de France participants. It turns out that Bernard Hinault comes at the first place, very closely followed by (*ex-aequo*) Eddy Merckx and Lance Armstrong; these three racers (robustly) dominate almost all other cyclists (a notable exception is Raymond Poulidor, who is dominated by no other cyclist).[7]

One possible criticism regarding our analysis is that it may be problematic to compare the performance of Tour de France racing cyclists in different eras. Indeed, cycling races have become increasingly internationalized. In addition, especially since the beginning of the nineties, racing cyclists have more and more become specialists in a single aspect of cycling, while in earlier times racing cyclists like Anquetil, Merckx and Hinault used to be outstanding not only in the Tour de France but also in many other important races. (Remark that this tendency to specialize actually also applies to the Tour itself; *e.g.*, we have nowadays specialists that focus on the final victory versus others that exclusively pursue the Green or the Red Polka Dot jersey.) In our application, we have chosen to put together racing cyclists of different eras in one and the same sample; in our opinion, this allows for most directly answering the question 'Who is the best Tour de France participant (since 1953)?'. Moreover, only seven racing cyclists who have been active in the last ten years are recorded in our list of the thirty highest ranked participants. Taking into account the importance of the Tour de France nowadays and the present-day's specialization, this seems to indicate that racing cyclists of earlier eras are not put at a disadvantage in our exercise. But if one still believes that different eras are intrinsically non-comparable, then one may conduct additional exercises that focus on subsamples (and that include, *e.g.*, only the racing cyclists who were able to score on at least one of the six dimensions in the last ten years). This essentially boils down to defining alternative (net-dominance) metrics (leading up to a potentially different overall ranking) from the matrix containing the pairwise dominance relationships (compare with the net-dominance metric results in Table 5 that basically summarize the pairwise dominance information in Table 4); and, in effect, it does not require additional computations regarding these pairwise dominance relationships.

[7] As a final qualification, we point out that these results are, of course, conditional upon the ordinal ranking of the different performance dimensions that is specified prior to the analysis. Some may argue that a different dimensional ordering is more appropriate. In such a case, sensitivity analysis is in order; the easy-to-compute nature of the proposed method makes such a sensivity check easy to implement in practice. As for our own application, we obtain that our results (*e.g.*, the identity of the top-three that dominates almost the entire pack of other racing cycists) are fairly robust with respect to alternative orderings that may be considered as 'reasonable'. For compactness, we did not include these results here, but they are available from the authors at simple request.

References

[1] Atkinson, A. and F. Bourguignon (1987), Income distribution and differences in needs, in G. Feiwel (ed.), *Arrow and the Foundations of the Theory of Economic Policy*, MacMillan, London.

[2] Cherchye, L. (2001), Using data envelopment analysis to assess macroeconomic policy performance, *Applied Economics*, 33, 407-416.

[3] Cherchye, L. and T. Kuosmanen (2004), Benchmarking sustainable development: a synthetic meta-index approach, forthcoming in M. McGillivray (ed.), *Perspectives on Inequality, Poverty and Human Well-Being*, United Nations University Press.

[4] Cherchye, L., W. Moesen and T. Van Puyenbroeck (2004), Legitimately diverse, yet comparable: synthesising social inclusion performance in the EU, *Journal of Common Market Studies*, to appear.

[5] Cooper, W.W., L. M. Seiford and K. Tone (2000), *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Kluwer Academic Publishers, Dordrecht.

[6] Coupé, T. (2003), Revealed performances: worldwide rankings of economists and economics departments 1990-2000, *Journal of the European Economic Association*, 1, 1309-1345.

[7] De Oliveira, E.H., and R. Callum (2004), Who's the best? Data Envelopment Analysis and ranking players in the National Football League, in P. M. Pardalos, S. Butenko and J. Gil-Lafuente (eds.), *Economics, Management and Optimization in Sports*, Springer Verlag.

[8] Einolf, K.W. (2004), Is winning everything? A Data Envelopment Analysis of Major League Baseball and the National Football League, *Journal of Sports Economics*, 5, 127-151.

[9] Evans, D., J. Lauer, C.A.K. Lovell, and C. Murray (2002), *World Health Systems Performance Revisited*, mimeo.

[10] Fukada-Parr, S. *et al.* (2004), *Human Development Report 2004: Cultural Liberty in Today's Diverse World*, United Nations Development Programme, New York.

[11] Haas, D., M.G. Kocher and M. Sutter (2004), Measuring the efficiency of German football teams by Data Envelopment Analysis, *Central European Journal of Operations Research*, to appear.

[12] Joint Research Centre (2004), *Composite Indicators: An information server on composite indicators*, European Commission http://farmweb.jrc.cec.eu.int/ci/.

[13] Kalaitzidakis, P., T. Mamuneas and T. Stengos (2003), Rankings of academic journals and institutions in economics, *Journal of the European Economic Association*, 1, 1346-1366.

[14] L'Equipe (2002), *100 Ans de Tour de France*, Issy les Moulineaux, L'Equipe.

[15] Lozano, S., G. Villa, F. Guerrero and P. Cortés (2002), Measuring the performance of nations at the Summer Olympics using Data Envelopment Analysis, *Journal of the Operational Research Society*, 53, 501-511.

[16] Lubrano, M., A. Kirman and L. Bauwens (2003), Rankings of economics departments in Europe: a statistical approach, *Journal of the European Economic Association*, 1, 1367-1401.

[17] Melyn, W. and W. Moesen (1991), Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available, Public Economics Research paper 17, CES, KU Leuven.

[18] Société du Tour de France (2004), http://www.letour.fr.

Racing cyclist	Yellow jersey	Green jersey
A	3	2
B	2	1
C	1	2
D	0	4

Table 1: A fictitious dataset

	A	B	C	D
A	1	0	0	0
B	1	1	0	0
C	1	1	1	0
D	1	0	0	1

Table 2: A fictitious dominance matrix

Note: Entry equal to 1 (0) implies that the racing cyclist in the associated row is (not) dominated by the racing cyclist in the associated column.

	Yellow jersey	Second place	Third place	Day victory	Green	Red Polka Dot
0	474	462	458	17	468	471
1	16	30	34	280	20	18
2	2	4	5	79	5	6
3	2	1	1	38	4	1
4	0	0	0	24	1	0
5	4	1	1	14	0	0
6	1	1	0	10	1	2
7	0	0	0	12	0	1
8	0	0	0	4	0	0
9	0	0	0	7	0	0
10	0	0	0	5	0	0
12	0	0	0	3	0	0
15	0	0	0	1	0	0
16	0	0	0	1	0	0
21	0	0	0	1	0	0
22	0	0	0	1	0	0
28	0	0	0	1	0	0
34	0	0	0	1	0	0

Table 3: Frequency table

	Armstrong	Hinault	Merckx	Anquetil	Indurain	LeMond	Bobet
Armstrong, Lance	1	0	0	0	0	0	0
Hinault, Bernard	0	1	0	0	0	0	0
Merckx, Eddy	0	0	1	0	0	0	0
Anquetil, Jacques	1	1	1	1	0	0	0
Indurain, Miguel	1	1	1	1	1	0	0
LeMond, Greg	1	1	1	1	1	1	0
Bobet, Louison	1	1	1	1	1	1	1
Fignon, Laurent	1	1	1	1	1	0	0
Thévenet, Bernard	1	1	1	1	1	0	0
Zoetemelk, Joop	0	1	0	0	0	0	0
Ullrich, Jan	1	1	1	0	0	0	0
Van Impe, Lucien	1	1	1	1	0	0	0
Bahamontès, Federico	1	1	1	1	1	0	0
Delgado, Pedro	1	1	1	1	1	1	1
Janssen, Jan	1	1	1	1	1	0	0
Gimondi, Felice	1	1	1	1	1	1	1
Pingeon, Roger	1	1	1	1	1	1	1
Gaul, Charly	1	1	1	1	1	0	0
Pantani, Marco	1	1	1	1	1	0	0
Riis, Bjarne	1	1	1	1	1	1	1
Roche, Stephen	1	1	1	1	1	1	1
Ocana, Luis	1	1	1	1	1	1	1
Nencini, Gastone	1	1	1	1	1	1	1
Aimar, Lucien	1	1	1	1	1	1	1
Walkowiak, Roger	1	1	1	1	1	1	1
Poulidor, Raymond	0	0	0	0	0	0	0
Virenque, Richard	1	1	1	1	1	0	0
Zabel, Erik	1	1	1	1	0	0	0

Table 4: Dominance matrix

Note: Entry equal to 1 (0) implies that the racing cyclist in the associated row is (not) dominated by the racing cyclist in the associated column.

Nr.	Racing cyclist	Yellow	Second	Third	Red Polka Dot	Green	Day	Dominating	Dominated	Net dominating
1	Hinault, Bernard	5	2	0	1	1	28	496	1	495
2	Armstrong, Lance	6	0	0	0	0	21	495	1	494
	Merckx, Eddy	5	1	0	2	3	34	495	1	494
4	Anquetil, Jacques	5	0	1	0	0	16	492	4	488
5	Zoetemelk, Joop	1	6	0	0	0	10	485	2	483
	Indurain, Miguel	5	0	0	0	0	12	488	5	483
7	Van Impe, Lucien	1	1	3	6	0	9	484	5	479
8	Ullrich, Jan	1	5	0	0	0	7	481	5	476
9	Thévenet, Bernard	2	1	0	0	0	9	482	7	475
	Fignon, Laurent	2	1	0	0	0	9	482	7	475
11	LeMond, Greg	3	1	1	0	0	5	477	6	471
12	Poulidor, Raymond	0	3	5	0	0	7	470	1	469
	Bahamontès, Federico	1	1	1	6	0	7	477	8	469
	Bobet, Louison	3	0	0	0	0	7	476	7	469
15	Gaul, Charly	1	0	2	2	0	10	474	8	466
16	Janssen, Jan	1	1	0	0	3	7	471	11	460
17	Pantani, Marco	1	0	2	0	0	8	471	12	459
18	Gimondi, Felice	1	1	0	0	0	7	468	15	453
19	Virenque, Richard	0	1	1	7	0	7	460	9	451
20	Ocana, Luis	1	0	0	0	0	9	463	16	447
	Delgado, Pedro	1	1	1	0	0	4	461	14	447
22	Pingeon, Roger	1	1	0	0	0	4	451	17	434
23	Chiappucci, Claudio	0	2	1	2	0	3	448	15	433
24	Riis, Bjarne	1	0	1	0	0	4	448	20	428
25	Van Springel, Herman	0	1	0	0	1	5	444	22	422
26	Jimenez, Julio	0	1	0	3	0	5	442	21	421
27	Bugno, Gianni	0	1	1	0	0	4	443	24	419
	Darrigade, André	0	0	0	0	2	22	423	4	419
29	Maertens, Freddy	0	0	0	0	3	15	422	6	416
30	Zabel, Erik	0	0	0	0	6	12	421	7	414

Table 5: A ranking of modern Tour de France racing cyclists