CENTRUM VOOR ECONOMISCHE STUDIEN

DISCUSSION PAPER

THE MEASUREMENT OF MACROECONOMIC PERFORMANCE: COMPARISON OF DEA-BASED ALTERNATIVES

by

Laurens CHERCHYE

Public Economics

Center for Economic Studies
Discussion Paper Series DPS 98.29

August 1998

Katholieke Universiteit Leuven
Departement Economie
Naamsestraat 69
B-3000 Leuven

The Measurement of Macroeconomic Performance: Comparison of DEA-Based Alternatives

Laurens Cherchye
Centrum voor Economische Studiën
KU Leuven
Naamsestraat 69
B-3000 Leuven, Belgium
Laurens.Cherchye@econ.kuleuven.ac.be
+32 16.32.68.54

August 1998

Abstract

It is common practice to summarise the macroeconomic performance of countries in terms of the four well-known dimensions captured by the magic diamond of the OECD. This paper provides a comparison of several synthetic indicators which merge the four separate indicators into one single statistic. These indicators are inspired by data envelopment analysis (DEA)-based models, which allow an unequal weighting of the different economic objectives. The calculated weights then act as proxies for the true policy priorities. Comparison of the models focuses on the underlying assumptions as well as on the empirical results they generate.

Keywords: macroeconomic performance, data envelopment analysis, unequal weighting

1. Introduction

Traditionally macroeconomic performance is measured as the extent to which policy makers reach their macroeconomic objectives. Policy objectives are then usually summarised in terms of the GDP growth rate, the inflation rate, the unemployment rate and the surplus or deficit on the current account of the balance of payments. The same four dimensions are captured by the magic diamond of the OECD [see e.g. Economic Outlook 41 (1987)].[1]

One might claim that also other dimensions should be incorporated in economic performance analysis. Asher, Defina and Thanawala (1993) argue that a full evaluation of economic performance must include measures of distributive justice. Lovell, Pastor and Turner (1995), on the other hand, suggest also to look at environmental indicators. By considering only four dimensions one assumes that the macroeconomic policy objectives of each country are solely defined in terms of these dimensions, which is indeed rather restrictive. This study does not aim to set out a more adequate set of policy objectives. Instead, the four traditional indicators are taken as a starting point for economic performance evaluation.

Though single indicators separately provide useful information, it remains difficult to rank countries in terms of economic performance. More often than not one country does not dominate another country in all four dimensions. As such, there is a need for synthetic measures, which merge several indicators into one single statistic. In the literature there have been some attempts to construct such a measure. A first one is the discomfort or misery-index [Mc Cracken et al. (1977)], originally developed by Okun and therefore also known as the *Okun index*. This measure is obtained as the sum of the unemployment rate and the inflation rate. Obviously, it anticipates the trade-off between inflation and unemployment. The interpretation is immediate: better performance is associated with lower values and vice versa. The *Calmfors-Driffill index*, developed by Calmfors and Driffill (1988), which adds the unemployment rate to the deficit on the current account, constitutes a second proposal. A lower value is again associated with better macroeconomic performance. In fact one penalises countries pursuing a low level of unemployment through an expansionary policy leading to a current account deficit.

Both the Okun and the Calmfors-Driffill index are obtained by adding two single indicators. The fact that only two and not four dimensions are considered is a first, minor inconvenience. More important is that the simple addition implies equal weighting of the economic objectives, which reflects the implicit assumption that policy makers find all policy goals equally important, or, in other words, that no priorities are attributed. This seems an unrealistic hypothesis. We would therefore like indicators which are obtained by unequally weighting the different macroeconomic objectives, with the weight of each objective being determined by the priority that is given to it. Data envelopment analysis (DEA) appears particularly well suited to obtain such synthetic indicators. The objective of this study consists in comparing different DEA-based measures theoretically (section 2) as well as empirically (section 3), comparing the results obtained for a sample of 20 OECD countries over the period 1992-1996. The main findings are reproduced in a final section.

2. DEA-based synthetic performance indicators

How could policy weights be determined when only assuming that they should sum up to one? One way is to look at what governments declare as being their policy objectives [Melyn en Moesen (1991)]. By examining the official government memoranda, which describe the stated policy objectives, weights could be derived on the basis of the place and the space occupied by each objective, hereby assuming that more important goals are to be found in the beginning and will be more developed.

This procedure is time consuming and requires a lot of information. Moreover, there is the more fundamental problem that publicly stated objectives do not always coincide with the true convictions of the policy makers. Stated preferences may deviate from revealed preferences. Therefore another line of approach is followed, where use is made of DEA models. DEA, first introduced by Charnes, Cooper and Rhodes (1978 and 1979), is a linear programming technique which is characterised by *benefit of the doubt weighting*. That is, if a country performs (relatively) well in a particular macroeconomic dimension, then a high relative weight is accorded to that dimension. Remember that synthetic indicators in fact aim at measuring the macroeconomic policy performance. As such, benefit of the doubt weighting reflects the assumption that good performance reflects high policy priority. This does not seem an unreasonable hypothesis, given the problems associated with the alternative.

Note, however, that DEA models ignore the possibility of asymmetric shocks [see Easterly, Kremer, Pritchett and Summers (1993)]. One should thus remain cautious in interpreting the computed weights and adequate information about exogenous shocks is desirable. On the other hand, DEA models could also constitute useful instruments in analysing the effects of asymmetric shocks. By studying a period before and after a particular shock or event, one could identify the effects it generated. An example is provided by Moesen and Cherchye (1997), who compared the performance of twenty OECD countries, half of which belonged to the European Union, in the quinquennial periods before and after the Maastricht Treaty so as to analyse the impact of pursuing the Maastricht convergence criteria on overall economic performance and policy priorities.

DEA was originally developed for operational research, in particular efficiency measurement. Performance evaluation as it is applied here deviates from efficiency evaluation in that only outputs are regarded and inputs are no longer relevant. One only considers the results and the way in which these are achieved is unimportant. In fact one measures efficiency with inputs put equal for all observations (in casu countries). This boils down to all reference to inputs being omitted in the original models.

The kind of models adressed in this paper must thus be distinguished from those used to analyse the macroeconomic efficiency (or productivity) of countries as in the study by Färe, Grosskopf, Norris and Zhang (1994), where DEA was used to construct a productivity index for 17 OECD countries. GDP was employed as aggregate output measure and capital stock and employment as aggregate input measures. Brockett, Golany and Li (1996) did a similar exercise, using the same input and output indicators. These authors studied the ratio between aggregate output and input. The difference with macroeconomic performance measurement is clear. Productivity studies examine the efficiency of the total economy, whereas studies about macroeconomic performance seek to identify those countries which achieve the best policy results, whether these results are attained in a (politically) efficient way or not.

Before starting with the study of the models, let me first introduce some notation. The total number of countries is denoted by n. The real GDP-growth rate will be represented by y, the unemployment rate by u, the inflation rate by p and the surplus or deficit on the current account as a percentage of GDP by f.

2.1. LIMEP 1

The first synthetic performance measure that will be treated is the LIMEP 1 [Melyn and Moesen (1991) and Moesen and Cherchye (1997)]. LIMEP stands for *L*euven *I*ndex of *M*acro*e*conomic *P*erformance. The model used is not really a DEA model, but it expresses the same idea of benefit of the doubt weighting. The construction of the LIMEP 1 proceeds in two steps. In a first step the original data are normalised. At the same time the economic 'bads' (unemployment and inflation) are converted into 'goods'. The values of the normalised indicators will vary between zero and one, zero corresponding to the worst performance in the sample and one corresponding to the best performance. Higher numerical values reflect better relative performance. The normalised counterparts of y, p, u

and f are denoted by respectively y^n, p^n, u^n and f^n. If subscript 'min' refers to the lowest indicator value over all countries in the sample and subscript 'max' to the highest indicator value, then for each country i the normalised values are computed as follows:

$$y_i^n = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \tag{2.1}$$

$$p_i^n = \frac{p_{\max} - p_i}{p_{\max} - p_{\min}} \tag{2.2}$$

$$u_i^n = \frac{u_{\max} - u_i}{u_{\max} - y_{\min}} \tag{2.3}$$

$$f_i^n = \frac{f_i - f_{\min}}{f_{\max} - f_{\min}} \tag{2.4}$$

The last step consists in according weights to the four macroeconomic dimensions. Here it is assumed that priorities are attributed. This idea is expressed by imposing that each objective should get a weight of either 0.10, 0.20, 0.30 or 0.40. The highest weight (0.40) should be attached to the dimension for which the country under consideration performs relatively best, the weight 0.30 should be given to the dimension for which relative performance is second best, and so on. The following model should then be solved for each country i (LIMEP 1):

$$\max_{a_{1i},\, a_{2i},\, a_{3i},\, a_{4i}} \quad \text{LIMEP } 1_i = a_{1i}\, y_i^n + a_{2i}\ p_i^n + a_{3i}\ u_i^n + a_{4i}\ f_i^n$$

s.t.

$$\begin{aligned} a_{1i} + a_{2i} + a_{3i} + a_{4i} &= 1 \\ a_{ji} + a_{ki} &\geq 0.30 \qquad j,k = 1,2,3,4 \quad j \neq k \\ 0.40 &\geq a_{ji} \geq 0.10 \qquad j = 1,2,3,4 \end{aligned}$$

The generosity of the model is immediate. The weights a_{1i}, a_{2i}, a_{3i} and a_{4i} are chosen in such a way that no other possible combination of weights would yield a higher value for the LIMEP 1. The program is solved for each country. The value of the LIMEP 1 will always be between zero and one.

The model is rather simple and very intuitive, which are two important qualities. The assumption that priorities should be attributed is more problematic. Some countries perform equally well in more than one dimension. E.g., one country may as well have the lowest unemployment rate as the highest growth rate. In this case one should in fact attach an equal weight of 0.35 to both indicators. This possibility is not excluded a priori by the formulation of the model. But whether 0.35 will be selected as the weight for the two indicators instead of the weights 0.30 and 0.40 depends on the algorithm that is followed.

Note that even when a clear ranking between the several dimensions can be made, there is no a priori reason why each weight should be 0.10, 0.20, 0.30 or 0.40. The model formulation, though, forces the weight 0.40 to be accorded to the indicator for which the country under consideration performs relatively best, the weight 0.30 to be attached to the indicator for which the country's relative performance is second best, and so on. The inequalities in the model de facto become equalities. In reality it could equally well be that for one country real growth gets a weight of, say, 0.15, unemployment a weight of 0.18, inflation a weight of 0.46 and, finally, the current account a weight of 0.21. The fact that this possibility is not captured by the model can be regarded as a weakness.

2.2. LIMEP 2

Melyn en Moesen (1991) also proposed another synthetic indicator, viz. the LIMEP 2. The LIMEP 2 model is based on the input oriented CCR model [Charnes, Cooper and Rhodes 1978 and 1979]. This model was originally constructed to provide a 'best practice envelope' in an input-output setting, assuming constant returns to scale. The latter is of course only an appropriate concept when one is working with production functions. If only outputs are considered, the assumption must be reinterpreted and this creates a fundamental problem, as will be shown below.

To calculate the LIMEP 2 one works with the original data. The economic 'bads', p and u, get a negative sign in the objective function. The weights still have to sum up to one, but are less restricted than before. The only restriction is that they should be at least 0.10, which creates a greater latitude. It seems intuitively clear that some lower bound should be respected by the weights.[2] It is indeed not very probable that countries attach an equal weight to each objective, but it is perhaps even less likely that one objective gets a weight of, say, 0.99 while all the other objectives together get a weight of only 0.01. The fact that weights can be chosen more freely appears to be an improvement over the model in the previous section. The linear programming problem, which should again be solved for each country i, can be written as (LIMEP 2):

$$\max_{b_{1i},\, b_{2i},\, b_{3i},\, b_{4i}} \quad \text{LIMEP } 2_i = b_{1i}\, y_i - b_{2i}\, p_i - b_{3i} u_i + b_{4i} f_i$$

s.t.

$$\begin{aligned} b_{1i}\, y_k - b_{2i}\, p_k - b_{3i} u_k + b_{4i} f_k &\leq 1 \qquad k = 1, ..., n && (2.5)\\ b_{1i} + b_{4i} + b_{3i} + b_{4i} &= 1 && (2.6)\\ b_{ji} &\geq 0.10 \qquad j = 1, 2, 3, 4 \end{aligned}$$

Intuition is straightforward. The fact that weights are computed so that the highest possible LIMEP 2 value is achieved for the country under consideration reflects the idea of benefit of the doubt weighting. By solving the model for each country we get a relative ranking. The best performers get a LIMEP 2 score of one. In general, the overall performance of an economy is better the closer the score gets to one. Remark that the value of the LIMEP 2 can be negative, for it is not bounded below. This has some consequences for the interpretation of the LIMEP 2 score, which have to do with the radial nature of the CCR model.

In an input-output setting, radial models seek for each observation the value which gives the equiproportionate expansion in all outputs (output orientation) or reduction in all inputs (input orientation), required to achieve parity with the best practice observed in the sample. Now, however, the LIMEP 2 value can be negative and it seems difficult to interpret the LIMEP 2 scores. Normally negative output values are not implemented directly into a radial model, and it appears that this is not an adequate procedure.

Suppose for the moment that this problem could be overcome by using alternatives y^*, p^*, u^* and f^* instead of y, p, u and f, with y^*, p^*, u^* and f^* only taking positive values and can all be interpreted as goods (so that we only have plus signs in the objective function). Note that y^* and f^* are not just translations of y and f, as radial models are not translation invariant (see e.g. Ali and Seiford (1990)). Using these new measures in the LIMEP 2 model, restriction (2.6) becomes problematic in combination with restriction (2.5). Suppose e.g. that $(y_i^*, p_i^*, u_i^*, f_i^*)$ equals $(2, 1, 3, 2)$. It is clear that in this case (2.5) and (2.6) cannot be satisfied simultaneously. On the other hand, if all y^*, p^*, u^* and f^* values were within the interval $(0, 1]$, then restriction (2.5) would no longer be binding because of (2.6) and the maximisation procedure would become trivial. Clearly, other constraints should be used to impose upper bounds on the weights. One could for example replace (2.6) by the following:

$$b_{ji} \geq \frac{1}{10}(b_{1i} + b_{2i} + b_{3i} + b_{4i}) \qquad j = 1, 2, 3, 4$$

By confining oneself to strictly positive data entries and by using appropriate weight restrictions the input oriented CCR model can be saved. But, as inputs are regarded to be irrelevant in the curerent setting, an output oriented model would constitute a better alternative. A fundamental problem then arises for the CCR model, of which the output oriented version looks as follows:[3]

$$\max_{\varphi_i,\lambda,s_{yi},s_{pi},s_{ui},s_{fi}} \quad z_i = \varphi_i + \varepsilon\ (s_{yi} + s_{pi} + s_{ui} + s_{fi})$$

s.t.

$$\begin{aligned} \varphi_i\ x_i^* - \sum_{k=1}^{n} \lambda_k\ x_k^* + s_{xi} &= 0 \qquad x = y, p, u, f \\ \lambda_k &\geq 0 \qquad k = 1, ..., n \\ s_{xi} &\geq 0 \qquad x = y, p, u, f \end{aligned}$$

The resulting output performance measure equals one for the best performers, and is above one for worse performers. Obviously, solving this model would always give a z_i value of ∞. The reason is that input restrictions are omitted from the original CCR model. How should this difficulty be coped with? One way is to add further restrictions on the λ's. An interesting choice is the convexity constraint $\sum_k \lambda_k = 1$, which yields the BCC model [Banker, Charnes and Cooper (1984)]. In an input-output setting the convexity constraint represents the assumption of variable returns to scale. In the current context this assumption should be interpreted as a variable 'trade-off' between the several outputs or macroeconomic services, which appears realistic. Remark that the possibility of a constant trade-off is not excluded. No a priori assumption about the nature of the trade-offs has to be made.

There still remains a final problem concerning the radial nature of the model, which follows from the presence of slacks. If these slacks are insignificantly small and do not systematically occur, this is only a minor problem. Lovell and Pastor (1995) found, however, when aplying the BCC model, that this is not the case for macroeconomic performance measurement. Slacks did not only appear frequently, they were also large in many instances. This has clear intuition: countries may perform relatively well in some macroeconomic policy dimensions while there is only a poor performane with respect to other dimensions. Obviously, in such cases it does seem hard to defend that a measure capturing the maximum equiproportionate expansion in all outputs will adequately reflect the overall economic policy performance. For this reason they proposed a non-radial model, which will be the subject of the next subsection.

2.3. GEM

The non-radial model proposed by Lovell and Pastor (1995) is the extended additive DEA model, introduced by Charnes, Cooper, Rousseau and Semple (1987). The same model was used by Lovell, Pastor and Turner (1995). Relative slacks are implemented directly into the objective function. As before the y, p, u and f are substituted for by appropriate alternatives y^*, p^*, u^* and f^*. The model which should be solved for each country i looks as follows (GEM-Primal):

$$\max_{\lambda,s_{yi},s_{pi},s_{ui},s_{fi}} \quad \Omega_i = \frac{s_{yi}}{y_i^*} + \frac{s_{pi}}{p_i^*} + \frac{s_{ui}}{u_i^*} + \frac{s_{fi}}{f_i^*} \tag{2.7}$$

s.t.

$$\begin{aligned} \sum_{k=1}^{n} \lambda_k\ x_k^* - s_{xi} &= x_i^* \qquad x = y, p, u, f \\ \sum_{k=1}^{n} \lambda_k &= 1 \end{aligned}$$

$$\begin{array}{rcll} \lambda_k & \geq & 0 & k = 1, ..., n \\ s_{xi} & \geq & 0 & x = y, p, u, f \end{array}$$

Each country which is not 'best practice' is compared to a Pareto efficient reference point. Obviously, the difficulty with the above problem is that the objective (2.7) lacks a natural interpretation as a relative efficiency measure. It is zero for an efficient observation and positive for an inefficient observation, but no real meaning attaches to the magnitude of a positive value, since the objective is unbounded above. Therefore Lovell and Pastor replaced (2.7) with a new objective, one which has a natural interpretation as a relative efficiency measure. This objective is given by:

$$\min_{\lambda, s_{yi}, s_{pi}, s_{ui}, s_{fi}} \text{GEM}_i = \left[1 + \frac{1}{4}\left(\frac{s_{yi}}{y_i^*} + \frac{s_{pi}}{p_i^*} + \frac{s_{ui}}{u_i^*} + \frac{s_{fi}}{f_i^*}\right)\right]^{-1}$$

GEM stands for *G*lobal *E*fficiency *M*easure. Clearly, the GEM can be obtained indirectly from the solution of the (linear) extended additive model above. The GEM value equals one if and only if the country under consideration is situated on the (Pareto efficient) best practice frontier or, in other words, if all slacks are zero. Values vary within the interval $]0, 1]$.

In the above model no lower bound is imposed on the relative weights attributed to the several macroeconomic objectives, while the possibility of importing restrictions on the policy weights seems desirable. In order to evaluate this possibility the dual formulation of the above model should be considered. After converting the obtained minimisation problem into a maximisation problem, we get (GEM-dual):

$$\max_{c_{1i},\, c_{2i},\, c_{3i},\, c_{4i}, \varpi_i} q_i = c_{1i}\, y_i^* + c_{2i}\;\, p_i^* + c_{3i}\, u_i^* + c_{4i}\, f_i^* + \varpi_i$$

s.t.

$$\begin{array}{rcll} c_{1i}\, y_k^* + c_{2i}\;\, p_k^* + c_{3i}\, u_k^* + c_{4i}\, f_k^* + \varpi_i & \leq & 0 & k = 1, ..., n \\ c_{ji} & \geq & \dfrac{1}{x_i^*} & j = 1, 2, 3, 4 \quad x = y, p, u, f \\ \varpi_i & = & free & \end{array}$$

Because of the duality theorem of linear programming, it holds that $\Omega_i^* = (-q_i^*)$, where the superscript $(^*)$ denotes an optimal value. The term ϖ_i is associated with the convexity constraint in the GEM-primal model and determines the intercept of the hyperplane represented by the objective function [Seiford and Thrall (1990)].

Clearly, lower bounds of the output weights will differ for different observations, which implies that one cannot impose particular lower bound weight values like in the LIMEP 2 model. This appears to be a drawback. In the next subsection another model will be proposed to construct a synthetic performance indicator, viz. the LIMEP 3, which does not suffer from this inconvenience.

2.4. LIMEP 3

From above we can derive two conditions that should be met by the DEA model used to compute a synthetic macroeconomic performance indicator. First, the model should be non-radial, for the presence of slacks in radial models appears to be problematic. Further, the model should allow for imposing restrictions on the policy weights. A DEA model which satisfies both conditions is the additive model, developed by Charnes, Cooper, Golany, Seiford and Stutz (1985). For each country i, the model can be formally expressed as:

$$\min_{\lambda, s_{yi}, s_{pi}, s_{ui}, s_{fi}} v_i = -s_{yi} - s_{pi} - s_{ui} - s_{fi} \tag{2.8}$$

s.t.

$$\begin{aligned}\sum_{k=1}^{n}\lambda_k\ x_k^n - s_{xi} &= x_i^n \qquad x = y,p,u,f\\ \sum_{k=1}^{n}\lambda_k &= 1\\ \lambda_k &\geq 0 \qquad k=1,...,n\\ s_{xi} &\geq 0 \qquad x = y,p,u,f\end{aligned}$$

(2.8) has an analogous interpretation as (2.7). The only difference is that the objective function (2.8) equals the sum of *absolute* slacks instead of relative slacks. This implies, in order to treat each dimension on equal footing, that the four original variables should be replaced by alternatives which are expressed in the same scale. Therefore use is made of the normalised variables y^n, p^n, u^n and f^n (cf. (2.1) to (2.4)).

Suppose one wants to eliminate large differences beween relative weights, assuming for example that in reality each policy objective gets a relative weight of at least ten percent. When weight restrictions are to be imposed the dual formulation of the additive model seems more suitable. It looks as follows for each country i (additive-dual):

$$\max_{g_{1i},\ g_{2i},\ g_{3i},\ g_{4i},\varpi_i} w_i = g_{1i}\ y_i^n + g_{2i}\ p_i^n + g_{3i}u_i^n + g_{4i}f_i^n + \varpi_i \tag{2.9}$$

s.t.

$$\begin{aligned}g_{1i}\ y_k^n + g_{2i}\ p_k^n + g_{3i}u_k^n + g_{4i}f_k^n + \varpi_i &\leq 0 \qquad k=1,..,n\\ g_{ji} &\geq 1 \qquad j=1,2,3,4\\ \varpi_i &= free\end{aligned}$$

Again, the duality theorem guarantees that $v_i^* = w_i^*$, with (*) denoting an optimal value. Ensuring that relative weights are not below ten percent (LIMEP 3):

$$\max_{d_{1i},\ d_{2i},\ d_{3i},\ d_{4i},\varpi_i} w_i^{'} = d_{1i}\ y_i^n + d_{2i}\ p_i^n + d_{3i}u_i^n + d_{4i}f_i^n + \varpi_i \tag{2.10}$$

s.t.

$$\begin{aligned}d_{1i}\ y_k^n + d_{2i}\ \ p_k^n + d_{3i}\ u_k^n + d_{4i}\ f_k^n + \varpi_i &\leq 0 \qquad k=1,...,i,...,n\\ d_{ji} &\geq 0.10 \qquad j=1,2,3,4\\ d_{ji} &\geq \frac{1}{10}(d_{1i}\ + d_{2i}\ + d_{3i} + d_{4i}) \qquad j=1,2,3,4\\ \varpi_i &= free\end{aligned}$$

Including lower bound weight values 0.10 instead of 1 only changes the scale of the objective function value. Because normalised values are used, the solution of the above model will always lie within the interval $[-\frac{2}{5}, 0]$. It will be zero for the best performers, and the more it is below zero the worse the overall economic achievement. The model produces the lowest possible objective function value for a country which performs worst in all macroeconomic dimensions if and only if another country in the sample performs best in all macroeconomic dimensions. This has clear intuition: a country performs worst in all dimensions, although another country, acting under similar circumstances (which is a priori assumed), proves that it is possible to perform best in all dimensions.

Clearly, a country cannot perform any worse (given the sample), and so the former gets the lowest possible score.

For interpretational convenience, we would like the values to lie between zero and one. This is of course easily obtained by normalising objective (2.10) in the following way:

$$\max_{d_{1i},\, d_{2i},\, d_{3i},\, d_{4i}, \varpi_i} \quad \text{LIMEP } 3_i = 1 + \frac{5}{2}(d_{1i}\, y_i^n + d_{2i}\;\, p_i^n + d_{3i}\; u_i^n + d_{4i}\;\, f_i^n + \varpi_i)$$

We get a model that incorporates some attractive properties. It captures the assumption that there is a variable trade-off between several macroeconomic objectives and it allows to implement assumptions about realistic ranges of policy weights. Moreover, the model is non-radial which deals with the empirical objections put forward against the use of radial DEA models. Finally, it produces performance scores between zero and one, with higher scores corresponding to better performance.

2.5. Discussion

Four models to construct an unequally weighted synthetic macroeconomic performance indicator have been discussed. The first three models were taken from the existing literature and the fourth one was an own proposal. The LIMEP 1 model captures the basic idea of DEA models as it is characterised by benefit of the doubt weighting. It concerns a simple model with clear intuition. The only inconvenience is that, to calculate LIMEP 1 scores, priorities should be attributed exogeneously, reflected by the condition that weights should take values of either 0.10, 0.20, 0.30 or 0.40.

The other three models are more sophisticated. The LIMEP 2 model is based on the *input*-oriented CCR model. This is already a first inconvenience, as we are measuring performance and inputs are regarded to be irrelevant. Moreover the LIMEP 2 values lack a clear interpretation as they can be negative. Finally, the model is radial and thus gives rise to slacks which appear to be problematic in macroeconomic performance measurement.

That's why the non-radial GEM and LIMEP 3 models have been introduced. The GEM model also captures the assumption of a variable trade-off between the several macroeconomic services. But it does not allow for imposing particular lower bound values on the policy weights. The LIMEP 3 model, which is also characterised by the assumption of a variable trade-off, does not suffer from this inconvenience. It produces LIMEP 3 scores, which, like the GEM scores, have the attractive property of lying between zero and one, with lower scores corresponding to worse performance.

In fact the different models reflect particular assumptions about the distribution of policy weights. Intuitively, one is inclined to think that the relative weights of the different components of the synthetic performance indicator should respect some lower bound, but it is not clear a priori which value this lower bound should take. Moreover, one could wonder whether also upper bounds would be appropriate. In order to solve these and related problems additional information will be needed. In a microeconomic environment market information can be used to set the range of potential weight values (see e.g. Charnes, Cooper, Huang and Sun (1990)), but in a macroeconomic context adequate information is not readily available.

As it is not immediately obvious which value the lower weight bounds should take, it can not be said a priori wether the GEM model is better suited than the LIMEP 3 model or vice versa. It would be interesting to compare the empirical results of both models with those of the less sophisticated but intuitive LIMEP 1 model. This will be done in the next section.

A final remark concerns the distinction between flow and stock variables. The objective of macroeconomic performance measurement is to evaluate policy performance. It is clear that the overall strength of an economy in a certain period does not only depend on the quality of macroeconomic management but also on the shape of the economy at the beginning of the period under study [see Lovell and Pastor (1995)].[4] Therefore it seems recommendable only to consider flow variables when evaluating policy performance. Real growth, the inflation rate and the surplus or deficit on the current account are clearly such variables. The unemployment rate is not, however. The relative change in the unemployment rate would constitute a better alternative. Nevertheless, in the next section a stock variable will be used to represent the employment dimension as this appears to be common practice in the literature. The latter does not prevent it from being a weakness, however, especially with respect to the interpretation of the relative weights as indicators of policy priorities.

3. Empirical Comparison

The LIMEP 1, the GEM and the LIMEP 3 procedures will be applied to evaluate a sample of 20 OECD countries. The data are presented in section 3.1. The sample covers the period 1992-1996. Because of the possibility of shifts in macroeconomic performances or policy priorities results will be computed for each separate year. These results will be discussed in section 3.2. The purpose is to check whether the different models, which reflect different perceptions about the distribution of policy priorities, involve divergent views on the relative macroeconomic performance of countries.

In fact, four models will be compared. Since it is a priori not obvious which value the lower bound restriction on the weights in the LIMEP 3 model should take, also LIMEP 3' results will be evaluated. The latter are obtained from the additive-dual model. Clearly there are no restrictions on the relative weights. As the objective (2.9) lacks a natural interpretation as a relative efficiency measure, it is, analogously to Lovell and Pastor (1995), replaced by:

$$\min_{g_{1i},\, g_{2i},\, g_{3i},\, g_{4i}, \varpi_i} \text{LIMEP } 3_i^{'} = \left[1 + \frac{1}{4}\left(g_{1i}\, y_i^n + g_{2i}\, p_i^n + g_{3i} u_i^n + g_{4i} f_i^n + \varpi_i\right)\right]^{-1}$$

The resulting measure only takes values within the interval $[\frac{1}{2}, 1]$, which is immediate from the objective (2.8), knowing that slacks can at most equal one. Clearly, high correlation could be expected between the LIMEP 3 and the LIMEP 3' results, as the LIMEP 3 model only differs from the LIMEP 3' model in that 'extreme' policy weight combinations are eliminated.

DEA models seem appropriate for computing performance scores as they allow for benefit of the doubt weighting. As such, it appears interesting not only to compare scores but also weights assigned to the several dimensions for each country. That's why the GEM-dual model will be used to compute the GEM results and not the GEM-primal model.

3.1. Data

As real growth, a surplus or deficit on the current account, the inflation rate and the unemployment rate cannot be employed in the GEM model, the variables y, p, u and f are substituted for by the alternative measures y^*, p^*, u^* and f^*, which only take positive values and can all be interpreted as goods. These variables are defined as follows for each year t:

$$\begin{aligned}
y_t^* &= \frac{GDP_t}{GDP_{t-1}}\ (= 1 + y_t), \text{ where } GDP_t = \text{ Real Gross Domestic Product in } t \\
p_t^* &= \frac{CPI_{t-1}}{CPI_t}\ (= \frac{1}{1+p_t}), \text{ where } CPI_t = \text{ Consumer Price Index in } t \\
u_t^* &= 1 - u_t = \text{Employment Rate} \\
f_t^* &= \frac{X_t}{M_t}, \text{ where } X_t = \text{ Value of Exports in } t \text{ and } M_t = \text{Value of Imports in } t
\end{aligned}$$

To calculate the LIMEP 1, the LIMEP 3 and the LIMEP 3' results normalised variables are used. To obtain these an analogous procedure to the one presented in (2.1) and (2.4) is followed. A summary of the data is presented in table 1.

Table 1: Original data, period 1992-1996 (average values)

Country	y^*	p^*	u^*	f^*
Australia	1.039	0.977	0.903	0.984
Austria	1.017	0.971	0.941	0.844
Belgium	1.012	0.978	0.877	1.055
Canada	1.022	0.986	0.896	1.135
Denmark	1.023	0.981	0.890	1.158
Finland	1.014	0.985	0.834	1.006
France	1.013	0.981	0.883	1.058
Germany	1.015	0.972	0.908	1.110
Ireland	1.062	0.978	0.862	1.320
Italy	1.020	0.957	0.910	1.113
Japan	1.014	0.993	0.972	1.378
Netherlands	1.022	0.975	0.933	1.095
New Zealand	1.032	0.980	0.919	1.030
Norway	1.038	0.981	0.944	1.324
Portugal	1.015	0.947	0.938	0.659
Spain	1.013	0.955	0.777	0.768
Sweden	1.009	0.975	0.926	1.195
Switzerland	0.999	0.979	0.959	1.021
United Kingdom	1.010	0.973	0.891	0.891
USA	1.026	0.972	0.937	0.789
Mean	1.021	0.975	0.914	1.054
Standard Deviation	0.016	0.011	0.050	0.225

Sources:
OECD, Economic Outlook, 1997
OECD, Main Economic Indicators, 1995 and 1997

3.2. Results

It was mentioned before that macroeconomic performance results should be interpreted with sufficient care. Nevertheless they provide useful information. For example they allow one to identify weak and strong performers. Moreover, the weights at least give some indication about the true policy priorities of governments. Finally, intertemporal comparison can help to distinguish performance and priority shifts. These points are illustrated by tables 2 and 3, which show results for Australia, Belgium, Germany, Ireland, Japan and the USA.[5,6]

Table 2 gives performance results for the year 1992. Results are presented for the four performance measures so as to get a first impression about the similarity of the results produced by the different models. All measures agree on Japan being the best performer. Rankings differ slightly. In general, however, bad performance for one measure also involves low ranking for another measure. There appears to be more disagreement about the weights, although some patterns can be disinguished. For instance, for Belgium high weights are always accorded to real growth and/or price stability. For Germany, Ireland and the USA growth would be by far the most important. As regards Japan, however, things are not that clear. But here it concerns a special case, as Japan achieves the highest possible score for the GEM, the LIMEP 3 and the LIMEP 3'. The maximisation procedure comes to an end as soon as it reaches the maximum score of one. Weight combinations different from the one produced but also resulting in a score of one could be possible. Which weight combination is selected then depends on the algorithm that is followed. The same remark holds for the Australian GEM, LIMEP 3 and LIMEP 3' weights. The possibility of multiple optimal solution and the associated difficulties in interpreting the weights only occur to a much smaller extent for countries not lying on the best practice frontier.

The LIMEP 3 results for all consecutive years are presented in table 3 for the same six countries. Japan remains on top for whole the period. Real objectives would be given the highest weights, but again many weight combinations are possible as Japan is assigned a score of one in each year. We also notice that Australia tumbles from a first place in the year 1992 to no worse than a seventeenth position in 1995. In 1996, then, performance gets a little better. Weights are hard to interpret for the year 1992 as a score of one is attained. In 1993 price stability appears to be preferred above the other objectives, whereas in the following years no priorities seem to be attributed. Next we find that Belgium experiences a dramatic relative performance deterioration in 1993 from which it doesn't manage to recover in the following years, except from a small recuperation in 1995. Also Germany performs only moderately, possibly a reflection of the unification. The German and the Belgian weight structures are the same, and each objective gets approximately the same policy weight for the whole period, except from the year 1992, where real growth is assigned a higher weight. This result somewhat runs against general intuition as the Maastricht convergence criteria promote a tight monetary policy so that we could expect low inflation to be politically more important. Ireland performs well, and gets from a fourth place in 1992 and 1993 to a first position in 1994, 1995 and 1996. In the first year real growth is accorded the highest weight, while in the second year also the current account result would get policy priority. In the following years the 'Emerald Tiger' gets a score of one which implies, as discussed before, that the computed weights can hardly be interpreted. The performance of the USA, finally, deters rather drastically over the period considered, falling back from a ninth place in 1992 to a seventeenth position in 1996. Slightly more concern goes to high growth, whereas the other objectives would be approximately equally important.

A final note applies with respect to the weights reported in table 3. In some instances, e.g. Ireland in 1992, highly unequal weights are obtained, which indicates that models allowing for unequal weighting of the different macroeconomic objectives may indeed reveal interesting information concerning policy priorities. On the other hand, in many cases we find equal weights for the four dimensions. One could conclude that unequal weighting procedures indeed create a surplus value but also that the a priori assumption that policy priorities will always be attributed seems hardly defendable. Equal weighting of the policy objectives should not be excluded. It further seems that the distribution of policy weights varies considerably over time and over countries.

To get a more precise idea about the similarity of the results for the four performance measures the correlation coefficients were computed. They are presented in tables 4 to 9. Table 4 gives the Pearson correlation between scores, which appears to be quite high. Note that the three lowest correlation coefficients are all associated with the GEM. One further notices that the LIMEP 3 has the highest correlation with each of the three alternatives. This result has clear intuition as the LIMEP 3 model takes an intermediate position: relative weights are indeed not bound to equal either ten, twenty, thirty or forty percent but they can also not be chosen completely freely as relative weight values below ten percent are impossible.

Additionally also the correlation between rankings has been computed. Results are reported in table 5. The Spearman correlation between the rankings appears to be very high (at least 95.5 percent). We conclude that the different models treated in section 3 lead to a large extent to the same conclusions regarding the performance of nations in terms of rankings and, to a somewhat lesser extent, in terms of scores.

The unequal weighting of their components constitutes the principal characteristic of the measures discussed here. Therefore it is interesting to consider the correlation between the weights used to determine the macroeconomic scores. From table 2 we could presume that, except from the correlation between the LIMEP 3 and the LIMEP 3' weights, this correlation will probably not be very high. This conjecture is affirmed by tables 6 to 9, which contain the relevant Pearson correlation coefficients. Especially the correlations between the f-weights is low. Looking at the correlation between weights associated with individual measures we see that the correlation between the weights used to construct the GEM and the LIMEP 3' measures, with the same underlying assumptions about policy priorities, is relatively high. Also the LIMEP 3 weights correlate (relatively) strongly with the weights associated with any of the other measures, which again results from the intermediate position which the LIMEP 3 model takes.

We conclude that, starting from the hypothesis that restricting economic policy objectives to four dimensions is not an inadequate procedure, other basic assumptions about the distribution of policy priorities can lead to significantly different weight combinations. The performance results are not so much affected. Not the calculated scores and certainly not the rankings. A priori assumptions about the allocation of policy weights over the four macroeconomic objectives, though they do not so much affect the ultimate scores and rankings, thus have a significant effect on the computed policy weights and so on the interpretation of policy performance. It seems important to pay attention to the weight restrictions implemented in the model used. Additional external information would constitute a great help.

Table 2: Results for Australia, Belgium, Germany, Ireland, Japan and the USA (year 1992)

	Score	*Ranking*	*y-weight*	*p-weight*	*u-weight*	*f-weight*
Australia:						
LIMEP 1	0.796	3	0.300	0.400	0.100	0.200
GEM	1.000	1	0.295	0.665	0.022	0.018
LIMEP 3	1.000	1	0.271	0.515	0.107	0.107
LIMEP 3'	1.000	1	0.090	0.730	0.090	0.090
Belgium:						
LIMEP 1	0.679	13	0.300	0.400	0.200	0.100
GEM	0.883	14	0.684	0.102	0.112	0.102
LIMEP 3	0.746	11	0.328	0.224	0.224	0.224
LIMEP 3'	0.797	11	0.328	0.224	0.224	0.224
Germany:						
LIMEP 1	0.676	14	0.400	0.200	0.300	0.100
GEM	0.911	10	0.678	0.109	0.113	0.099
LIMEP 3	0.779	8	0.328	0.224	0.224	0.224
LIMEP 3'	0.819	8	0.328	0.224	0.224	0.224
Ireland:						
LIMEP 1	0.795	4	0.400	0.200	0.100	0.300
GEM	1.000	1	0.899	0.035	0.040	0.027
LIMEP 3	0.968	4	0.700	0.100	0.100	0.100
LIMEP 3'	1.000	1	0.737	0.088	0.088	0.088
Japan:						
LIMEP 1	0.941	1	0.100	0.200	0.300*	0.400*
GEM	1.000	1	0.654	0.129	0.130	0.087
LIMEP 3	1.000	1	0.328	0.224	0.224	0.224
LIMEP 3'	1.000	1	0.090	0.731	0.090	0.090
USA:						
LIMEP1	0.713	8	0.400	0.300	0.200	0.100
GEM	0.857	15	0.697	0.095	0.100	0.109
LIMEP 3	0.776	9	0.328	0.224	0.224	0.224
LIMEP 3'	0.817	9	0.328	0.224	0.224	0.224

Table 3: **Results for Australia, Belgium, Germany, Ireland, Japan and the USA (LIMEP 3)**

	Score	*Ranking*	*y-weight*	*p-weight*	*u-weight*	*f-weight*
Australia:						
1992	1.000	1	0.271	0.515	0.107	0.107
1993	0.881	5	0.293	0.236	0.236	0.236
1994	0.862	6	0.418	0.194	0.194	0.194
1995	0.616	17	0.251	0.250	0.250	0.250
1996	0.694	10	0.250	0.250	0.250	0.250
Belgium:						
1992	0.746	11	0.328	0.224	0.224	0.224
1993	0.614	17	0.250	0.250	0.250	0.250
1994	0.694	17	0.250	0.250	0.250	0.250
1995	0.731	10	0.251	0.250	0.250	0.250
1996	0.617	16	0.250	0.250	0.250	0.250
Germany:						
1992	0.779	8	0.328	0.224	0.224	0.224
1993	0.618	16	0.250	0.250	0.250	0.250
1994	0.740	13	0.250	0.250	0.250	0.250
1995	0.768	7	0.251	0.250	0.250	0.250
1996	0.522	18	0.250	0.250	0.250	0.250
Ireland:						
1992	0.968	4	0.700	0.100	0.100	0.100
1993	0.955	4	0.361	0.100	0.100	0.439
1994	1.000	1	0.422	0.378	0.100	0.100
1995	1.000	1	0.251	0.250	0.250	0.250
1996	1.000	1	0.238	0.287	0.238	0.238
Japan:						
1992	1.000	1	0.328	0.224	0.224	0.224
1993	1.000	1	0.293	0.236	0.236	0.236
1994	1.000	1	0.201	0.201	0.201	0.397
1995	1.000	1	0.251	0.250	0.250	0.250
1996	1.000	1	0.238	0.287	0.238	0.238
USA:						
1992	0.776	9	0.328	0.224	0.224	0.224
1993	0.744	9	0.293	0.236	0.236	0.236
1994	0.723	15	0.250	0.250	0.250	0.250
1995	0.643	16	0.251	0.250	0.250	0.250
1996	0.589	17	0.250	0.250	0.250	0.250

Table 4 : Correlation between scores (Pearson)

	LIMEP 1	*GEM*	*LIMEP 3*	*LIMEP 3'*
LIMEP 1	1.000	0.797	0.949	0.917
GEM	0.797	1.000	0.886	0.873
LIMEP 3	0.949	0.886	1.000	0.980
LIMEP 3'	0.917	0.873	0.980	1.000

Table 5 : Correlation between rankings (Spearman)

	LIMEP 1	*GEM*	*LIMEP 3*	*LIMEP 3'*
LIMEP 1	1.000	0.994	0.998	0.998
GEM	0.994	1.000	0.995	0.955
LIMEP 3	0.998	0.995	1.000	0.999
LIMEP 3'	0.998	0.955	0.999	1.000

Table 6 : Correlation between y-weights (Pearson)

	LIMEP 1	*GEM*	*LIMEP 3*	*LIMEP 3'*
LIMEP 1	1.000	0.327	0.443	0.339
GEM	0.327	1.000	0.555	0.487
LIMEP 3	0.443	0.555	1.000	0.842
LIMEP 3'	0.339	0.487	0.842	1.000

Table 7 : Correlation between p-weights (Pearson)

	LIMEP 1	*GEM*	*LIMEP 3*	*LIMEP 3'*
LIMEP 1	1.000	0.275	0.255	0.186
GEM	0.275	1.000	0.665	0.560
LIMEP 3	0.255	0.665	1.000	0.865
LIMEP 3'	0.186	0.560	0.865	1.000

Table 8 : Correlation between u-weights (Pearson)

	LIMEP 1	*GEM*	*LIMEP 3*	*LIMEP 3'*
LIMEP 1	1.000	0.253	0.423	0.379
GEM	0.253	1.000	0.555	0.620
LIMEP 3	0.423	0.555	1.000	0.842
LIMEP 3'	0.379	0.620	0.842	1.000

Table 9 : Correlation between f-weights (Pearson)

	LIMEP 1	*GEM*	*LIMEP 3*	*LIMEP 3'*
LIMEP 1	1.000	0.086	0.155	0.059
GEM	0.086	1.000	0.266	0.452
LIMEP 3	0.155	0.266	1.000	0.630
LIMEP 3'	0.059	0.452	0.630	1.000

4. Summary and conclusions

In this paper several DEA-based performance indicators have been compared. Such indicators bring together several macroeconomic policy dimensions into one single statistic. It is common practice to summarise economic policy objectives in terms of the four dimensions captured by the magic diamond of the OECD. The fact that only four dimensions are considered appears restrictive. Though the aim of this study was not to compose a more appropriate set of performance dimensions, it is interesting to note here that the alternatives considered in sections 3 and 4 allow to incorporate also other dimensions.

DEA allows to unequally weight the different policy objectives, with each weight reflecting the degree of priority. In fact, benefit of the doubt weighting is applied. Good performance is associated with high policy priority and therefore implies high weighting in the performance measure. It is implicitly assumed that all countries are subject to the same (symmetric) shocks, a somewhat artificial assumption. Indeed, adequate information about asymmetric shocks during the period under study seems desirable in order to analyse performance results in an appropriate way.

Three proposals to construct a synthetic performance measure based on DEA have appeared in the literature, viz. the LIMEP 1, the LIMEP 2 and the GEM. The discussion of these three measures, together with a fourth and a fifth one, the LIMEP 3 and the LIMEP 3', formed the main subject of this study. There appear to be a number of conceptual problems associated with the LIMEP 2. For the other measures it is not a priori clear which one is best suited for macroeconomic performance measurement, though the underlying models reflect rather different assumptions about the distribution of policy priorities. The main reason is that additional information which allows to have a clear idea about reasonable sets of policy weights is generally lacking.

Empirical results for the different alternatives were compared so as to identify whether different assumptions about policy priorities have a substantial effect on the obtained scores, rankings and weights. We found that with respect to scores and rankings the impact was small and sometimes even negligible. The calculated weights, on the other hand, appear to be rather heavily affected.

Synthetic indicators can indeed provide an interesting contribution to macroeconomic policy evaluation. One should put the performance results into the right perspective, though. As it seems impossible to capture every macroeconomic policy objective it could perhaps be interesting to check the sensitivity of performance results with respect to different sets of single indicators. Further, external information allowing to identify a priori realistic ranges of possible weights for each objective and perhaps even for each country and each time period seems highly relevant. If adequate information could be gathered the LIMEP 3 procedure seems promising as it allows to impose all kinds of weight restrictions.

Acknowledgements

The author wishes to thank W. Moesen and T. Van Puyenbroeck for constructive comments. The usual disclaimer applies.

Notes

1. Inclusion of the current account may be subject to some controversy. Tinbergen (1952) argues that it does not so much concern an element of well-being as well a technical expression of a 'sound' policy. It could be argued that the fact that a surplus is associated with good macroeconomic performance originates from the European mercantilist nature. Nevertheless the current account is rather generally used as an indicator of macroeconomic performance.

2. Note that the value of 0.10 is arbitrarily chosen. It could equally well be 0.05, 0.15 or any other value as long as it is not above $1/d$, where d is the number of single indicators used in the construction of the synthetic performance measure.

3. Lower bound constraints imposed on the weights are omitted in order to simplify things. This does not affect the core of the argument.

4. In fact, Lovell and Pastor (1995) proposed to address this issue by normalising the four economic service indicators in such a way that each country starts out on an equal footing at the beginning of the period under study.

5. Note that the use of a stock variable as an indicator for employment performance somewhat weakens this argument.

6. Remark that the weights produced by the GEM, the LIMEP 3 and the LIMEP 3' procedures do not necessarily add up to one for each observation. As we want to compare the results of the four performance measures, however, the latter is a desired property. It is obtained by dividing the weights for each observation by their sum.

7. Note that both the GEM and the LIMEP 3' models allow the weights to be chosen freely but that correlation between the calculated weights is far from perfect, which indicates that also the model and the type of data that is used (original or transformed) have an impact.

References

Ali, A. and L.M. Seiford (1990), Translation Invariance in Data Envelopment Analysis, *Operations Research Letters* 9, 403-405.

Asher,M., R. Defina and K. Thanawalla (1993), The Misery Index: Only Part of the Story, *Challenge*, March-April 1993, 58-62.

Brockett, P.L., B. Golany and S. Li (1996), Analysis of Intertemporal Efficiency Trends Using Rank Statistics With an Application Evaluating the Macro Economic Performance of OECD Nations, *Mimeo*.

Banker, R.D., A. Charnes and W.W. Cooper (1984), Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management Science 30 (9)*, 1078-1092.

Calmfors, L. and J. Driffil (1988), Bargaining Structure, Corporatism and Macroeconomic Performance, *Economic Policy* 6, 13-61.

Charnes, A., W.W. Cooper, Z.M. Huang and D.B. Sun (1990), Polyhedral Cone-Ratio DEA Models with an Illustrative Application to Large Commercial Banks, *Journal of Econometrics* 46, 73-91.

Charnes, A., W.W. Cooper, B. Golany, L. Seiford and J. Stutz (1985), Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions, *Journal of Econometrics* 30 (1/2), 91-107.

Charnes, A., W.W. Cooper and E. Rhodes (1978), Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research* 2, 429-444.

Charnes, A., W.W. Cooper and E. Rhodes (1979), Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research* 3, 239.

Charnes, A., W.W. Cooper, J. Rousseau and J. Semple (1987), Data Envelopment Analysis and Axiomatic Notions of Efficiency and Reference Sets, *Research reports CCS* 558 (Center For Cybernetic Studies, Austin).

Easterly, W., M. Kremer, L. Pritchett and L. Summers (1993), Good Policy or Good Luck? Country Growth Performance and Temporary Shocks, *Journal of Monetary Economics*, December, 459-483.

Färe, R., S. Grosskopf, M. Norris and Z. Zhang (1994), Productivity Growth, Technical Progress and Efficiency Change in Industrialised Countries, *American Economic Review* 84 (1), 66-83.

Lovell, C.A.K., J.T. Pastor and J.A. Turner (1995), Measuring Macroeconomic Performance in the OECD: A Comparison of European and Non-European Countries, *European Journal of Operational Research*, forthcoming.

Lovell, C.A.K. and J.T. Pastor (1995), Macroeconomic Performance of Sixteen Ibero-American Countries Over the Period 1980-1991, *Mimeo*.

Mc Cracken, P. et al. (1977), *Towards Full Employment and Price Stability* (OECD, Paris).

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 (Centre for Economic Studies, Leuven).

Moesen, W. and Cherchye, L., 1997, The Macroeconomic Performance of Nations. Measurement and Perception, *Public Economics Research Paper* (Centre for Economic Studies, Leuven).

OECD (1987), *Economic Outlook 41* (OECD, Paris).

OECD (1997), *Economic Outlook 61* (OECD, Paris).

OECD (1995), *Main Economic Indicators 1995* (OECD, Paris).

OECD (1997), *Main Economic Indicators 1997* (OECD, Paris).

Seiford, L.M. and R.M. Thrall (1990), Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis, *Journal of Econometrics* 46 (1/2), 7-38.

Tinbergen, J. (1952), *On the Theory of Economic Policy. Contributions of Economic Analysis* (North-Holland Publishing Company, Amsterdam).

DISCUSSION PAPERS 1997

DPS 97.01 Hans Dewachter, Geert Gielens and Dirk Veestraeten, *An Assessment of Central Banks'. Stand on Exchange Rate Stabilization Policies*, February. (International Economics).

DPS 97.02 Anneleen Peeters, *How Hiring and Firing Costs Affect Labour Demand in a Model of Uncertainty*, February. (International Economics).

DPS 97.03 Hans Eyssen, *Are West-African Immigrants Discriminated in Cote D'Ivoire?*, March. (Development Economics).

DPS 97.04 Wim Lagae, *The Absorption of the Effects of Debt Relief Operations by European Export Credit Agencies: An Institutional Analysis*, March. (Development Economics).

DPS 97.05 Jenke ter Horst and Marno Verbeek, *Estimating Short-Run Persistence in Mutual Fund Performance*, March. (Econometrics).

DPS 97.06 Hans Dewachter and Hanno Lustig, *A Cross-Country Comparison of CPI as a Measure of Inflation*, April. (Financial Economics).

DPS 97.07 Yunus Aksoy and Yohanes E. Riyanto, *Exchange Rate Pass-Through in Vertically Related Markets*, May. (International Economics).

DPS 97.08 Paul De Grauwe, *The Indeterminacy of the Euro Conversion Rates. Why it Matters and how it can be solved*, June. (International Economics).

DPS 97.09 Jozef Konings, Hylke Vandenbussche and Reinhilde Veugelers, *Union Wage Bargaining and European Antidumping Policy*, June (Financial Economics).

DPS 97.10 Francis Vella and Marno Verbeek, *Using Rank Order as an Instrumental Variable: An Application to the Return to Schooling*, May. (Econometrics).

DPS 97.11 Jozef Konings and Patrick P. Walsh, *The Effect of Real Exchange Rate Movements and Ownership on the Life Cycle of Manufacturing Plants in Ireland (1973-1994)*, June. (Financial Economics).

DPS 97.12 Johan Eyckmans, *Balancedness of games with multilateral environmental externalities*, August. (Public Economics).

DPS 97.13 Patrick Van Cayseele and Dave FURTH, *Price Leadership and Buyouts*, September. (Financial Economics).

DPS 97.14 Mark De Broeck, Dominique Guillaume and Emmanuel Van der Stichele, *Small and Big Bangs in Bond Markets*, September (Financial Economics).

DPS 97.15 Hanno Lustig, *Re-examining the Synchronization of European Business Cycles*, September (Financial Economics).

DPS 97.16 João P. Cocco, Francisco J. Gomes and Pascal J. Maenhout, *A Two-Period Model of Consumption and Portfolio Choice With Incomplete Markets*, June (Financial Economics).

DPS 97.17 Stefan Dercon and Daniel Ayalew, *Demobilisation and Reintegration of Ex-Soldiers in Ehtiopia*, October (Development Economics).

DPS 97.18 Stefan Dercon and Pramila Krishnan, *In Sickness and in Health…Risk-Sharing within Households in Rural Ethiopia*, October (Development Economics).

DPS 97.19 Paul Collier, Stefan Dercon and John Mackinnon, *Density versus Quality in Health Care Provision: The Use of Household Data for Budgetary Choices in Ehtiopia* October (Development Economics).

DPS 97.20 van der Lijn Nick and Marno Verbeek, *Excess demand, repressed inflation, and forced saving in the Soviet Union,* October (Econometrics).

DPS 97.21 Lorelei Crisologo-Mendoza and Dirk Van de gaer, *Population Growth and Customary Law on Land: The Case of Cordillera Villages in the Philippines,* October (Development Economics).

DPS 97.22 Tom Van Puyenbroeck, *Two-stage optimal control problems subject to an isoperimetric constraint,* October (Public Economics).

DPS 97.23 Erik Schokkaert, Geert Dhaene and Carine Van de Voorde, *Risk Adjustment and the Trade-off Between Efficiency and Risk Selection*, November (Public Economics).

DPS 97.24 Paul De Grauwe, Hans Dewachter and Dirk Veestraeten, *Stochastic Process Switching and Stage III of EMU*, November (International Economics).

DPS 97.25 Dirk Heremans, *Regulation of Banking and Financial Intermediation,* November (Financial Economics).

DPS 97.26 Knud J. Munk, *Agricultural Policy a Public Economic Explanation,* November (Public Economics).

DPS 97.27 Hans Dewachter, *Can Markov Switching Models Replicate Chartist Profits in the Foreign Exchange Market?* November (International Economics).

DPS 97.28 Paul De Grauwe and Frauke Skudelny, *The Impact of EMU on Trade Flows,* December (International Economics).

DISCUSSION PAPERS 1998

DPS 98.01 Louis Baeck, *Thematisation and Canon Building in Post-War Development Studies*, January (Development Economics).

DPS 98.02 Hans Dewachter and Hanno Lustig, *Sticky Prices and the Nominal Effects of Real Shocks*, January (International Economics).

DPS 98.03 Ilse Frederickx, *Health in Rural Tanzania: The Determinants of Health Status, Health Care Demand and Health Care Choice,* January (Development Economics).

DPS 98.04 Paul De Grauwe, *European Unemployment. A Tale of Demand and Supply,* February (International Economics).

DPS 98.05 João Cocco, Francisco Gomes and Pascal Maenhout, *Consumption and Portfolio Choice over the Life-Cycle,* March (Financial Economics).

DPS 98.06 Yunus Aksoy and Hanno Lustig, *Circular Aspects of Exchange Rates and Market Structure,* January (International Economics).

DPS 98.07 André Decoster and Frederic Vermeulen, *Evaluation of the Empirical Performance of Two-Stage Budgeting AIDS, QUAIDS and Rotterdam Models Based on Weak Separability,* April (Public Economics).

DPS 98.08 Erik Schokkaert and Luc Van Ootegem, *Preference Variation and Private Donations,* April (Public Economics).

DPS 98.09 Erik Schokkaert, *Mr. Fairmind is Post-Welfarist: Opinions on Distributive Justice,* April (Public Economics).

DPS 98.10 Dirk Van de gaer, Michel Martinez and Erik Schokkaert, *Measuring Intergenerational Mobility and Equality of Opportunity*, April (Public Economics).

DPS 98.11 Paulo Augusto Nunes, *Testing the Validity of WTP values from a Contingent Valuation Survey in Portugal*, April (Public Economics).

DPS 98.12 Paulo Augusto Nunes, *Measuring the WTP for Recreation and Biodiversity Protection Programs,* April (Public Economics).

DPS 98.13 Laurens Cherchye and Tom Van Puyenbroeck, *Learning from Input-Output Mixes in DEA: A Proportional Measure for Slack-Based Efficient Projections,* February (Public Economics).

DPS 98.14 Jerzy Mycielski and Yohanes Riyanto, *On the Interaction between Taste and Distance and Its Implications on the Vertical Distribution Arrangement*, May (Financial Economics).

DPS 98.15 Jerzy Mycielski, Yohanes Riyanto and Filip Wuyts, *Product Differentiation and the Equilibrium Structure of the Manufacturer-Retailer Relationship,* May (Financial Economics).

DPS 98.16 Hans Degryse and Patrick Van Cayseele, *Relationship Lending within a Bank-based System: Evidence from European Small Business Data,* April (Financial Economics).

DPS 98.17 Pramila Krishnan, Tesfaye Gebre Selassie and Stefan Dercon, *The Urban Labour Market During Structural Adjustment: Ethiopia 1990-1997*, April (Development Economics).

DPS 98.18 Bart Capéau and Stefan Dercon, *Prices, Local Measurement Units and Subsistence Consumption in Rural Surveys: An Econometric Approach with an Application to Ethiopia*, March (Development Economics).

DPS 98.19 Stefan Dercon and Pramila Krishnan, *Changes in Poverty in Rural Ethiopia 1989-1995: Measurement, Robustness Tests and Decomposition*, March (Development Economics).

DPS 98.20 Jenke R. ter Horst, Theo E. Nijman and Marno Verbeek, *Eliminating Biases in Evaluating Mutual Fund Performance from a Survivorship Free Sample,* June (Econometrics).

DPS 98.21 Hilke Vandenbussche and Jozef Konings, *Globalization and the effects of national versus international competition on the labour market. Theory and evidence from Belgian firm level data,* August (Financial Economics).

DPS 98.22 Wim Moesen and Laurens Cherchye, *The Macroeconomic Performance of Nations Measurement and Perception,* August (Public Economics).

DPS 98.23 Anneleen Peeters, *Interim Employment and a Leading Indicator for the Belgian Labour Market*, September (International Economics, IERP 137).

DPS 98.24 Wessel Marquering and Marno Verbeek, *An Empirical Analysis of Intertemporal Asset Pricing Models with Transaction Costs and Habit Persistence,* September (Econometrics).

DPS 98.25 Filip Abraham and Joeri Van Rompuy, *Is Belgium ready for EMU? A look at national, sectoral and regional developments,* September (International Economics, IERP 138).

DPS 98.26 Sara Ochelen, Stef Proost and Kurt Van Dender, *Optimal Pricing for Urban Road Transport Externalities,* September (Public Economics).

DPS 98.27 Knud Munk, *Optimal Support to Low-skilled Households*, July (Public Economics).

DPS 98.28 Wim Moesen and Philippe Van Cauwenberge, *The Status of the Budget Constraint, Federalism and the Relative Size of Government: A Bureaucracy Approach,* September (Public Economics).

DPS 98.29 Laurens Cherchye, *The Measurement of Macroeconomic Performance: Comparison of DEA-Based Alternatives,* August (Public Economics).

DPS 98.30 Jürgen Janssens, *Volatility and Risk Premia on Belgian Secondary Long Term Government Bond Markets,* October (Financial Economics).

DPS 98.31 Stef Proost and Kurt Van Dender, *Effectiveness and Welfare Impacts of Alternative Policies to Address Atmospheric Pollution in Urban Road Transport* (Public Economics).

DPS 98.32 Inge Mayeres and Stef Proost, *Marginal Tax Reform, Externalities and Income Distribution*, (Public Economics).

DPS 98.33 André Decoster and Guy Van Camp, *The unit of analysis in microsimulation models for personal income taxes: fiscal unit or household?* (Public Economics).

DPS 98.34 Paul De Grauwe, Hans Dewachter and Yunus Aksoy, *The European Central Bank: Decision Rules and Macroeconomic Performance*, (International Economics).

DPS 98.35 Laurent Franckx, *The use of ambient inspections in environmental monitoring and enforcement when the inspection agency cannot commit itself to announced inspection probabilities*, (Public Economics).