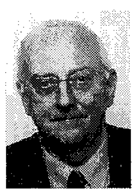
Tijdschrift voor Economie en Management
Vol. XLVII, 3, 2002

Measuring and Explaining the Productive Efficiency of Tax Offices: a Non-Parametric Best Practice Frontier Approach

door W. MOESEN en A. PERSOON*



Moesen Wim
Department of Economics,
Center for Economic Studies,
KULeuven, Leuven.

Persoons Annick
Department of Economics,
Center for Economic Studies,
KULeuven, Leuven.

ABSTRACT

In this paper we mimic an engineering approach to the "production" of tax offices. Essentially one dominant physical input (labour) is converted into heterogeneous non-monetary outputs such as the number of audited returns with a different degree of complexity.

Productive efficiency is evaluated against a best practice frontier using the non-parametric Free Disposal Hull (FDH) method and Data Envelopment Analysis (DEA). We first calculate efficiency measures for 289 regional tax offices, responsible for the personal income tax in Belgium. Next we explain the differences in efficiency scores in terms of characteristics related to managerial skills/culture and organizational structures.

* Financial support from the Belgian National Science Foundation (FDWTCA-PE/01/001) is gratefully acknowledged. Thanks also to the Ministry of Finance for providing the data and Lucy Amez and Tom Van Puyenbroeck for computational help.

I. INTRODUCTION

Fiscal stress, both at the level of national and local governments, has spurred a renewed interest among economists in the measurement of the productive efficiency of public administrations. Public agencies typically supply multiple outputs with heterogeneous inputs. Quite often the exact knowledge of the underlying production technology is rather weak. The construction of a non-parametric production frontier may then offer a solution.

Since the seminal paper of Farrell (1957) a substantial strand of axiomatic literature has treated the measurement of productive efficiency (see e.g. Färe, Groskopf and Lovell (1985)). An illuminating comparison of non-parametric with parametric estimation of technical efficiency in service production is provided in Bjurek, Hjalmarsson and Försund (1990). It is interesting to note that the practical tools for efficiency measurement have been supplied by the field of operational research. Applying linear programming techniques Charnes, Cooper and Rhodes ((1978), (1981)) have developed a method, which is widely known as Data Envelopment Analysis (DEA). In recent years Tulkens et al. ((1986), (1988)) have complemented this approach with the Free Disposal Hull (FDH) method. Common to both methods is the ambition to construct a production surface, which encompasses the actually observed best practices as close as possible. The degree of efficiency of any particular observation is then measured as the radial distance from this best practice frontier.

At the empirical level these methods (primarily DEA) have been extensively applied for the estimation of the efficiency scores of individual production units (cross-section) in various areas such as: health services, education, public transportation, post offices, municipalities, banking, insurance, ... (for an overview of empirical work both in the public and private sector, see Gulledge and Lovell (1992)).

On the contrary, at the revenue side of the government sector, empirical studies on the productive efficiency of tax offices seem to be rather rare. Although efficiency of the tax administration is one of the four "canons" recommended by Adam Smith (1776) in his long-standing treatise on taxation. Moreover, one notices that economists are more interested in the empirical investigation of individual compliance costs of taxation rather than the measurement of the operational efficiency of existing tax administrations (see Sandford, Godwin and Hardwick (1989); Slemrod (1992). Apparently, DEA and FDH have not yet been applied for an investigation of the relative efficiency of individual tax offices.

The purpose of this study is to measure and explain the differences in productive efficiency of individual tax offices. The sample covers the totality of regional tax offices, belonging to the Finance Ministry in Belgium and responsible for the personal income tax. On average 43 pct. of total national tax revenue is raised by the personal income tax, which makes it the most important single form of taxation in Belgium. These offices audit the tax returns, perform controls in loco, calculate the increases in the tax base for the underdeclarations, impose fines.... For these multiple outputs they employ one dominant input, which is labour.

Both the DEA and FDH method will be applied for the estimation and interpretation of the efficiency scores. To control for the sensitivity of the estimates we add an exercise to eliminate the outliers following a procedure outlined in Belsley, Kuh and Welsch (1980). As a further step we also want to explain the differences in efficiency taking into account scale, managerial and organizational characteristics. There we use the Tobit censored regression model.

The outline of the paper is as follows. The FDH and DEA methods are presented in section II. The data sources are discussed in section III. In the next section IV the efficiency results are reported and interpreted. Section V analyses the sensitivity of the results. An explanation of the efficiency indicators is given in section VI. Section VII draws some conclusions.

II. METHODOLOGY

Consider a tax office as a production unit or, in general terms as a decision making unit (DMU). Productive efficiency is achieved by a DMU if an increase in any output requires a reduction in at least one other output or an increase in at least one input; and if a reduction in any input requires an increase in at least one other input or a reduction in at least one output (Koopmans (1951)). Thus efficiency is represented by the attainment of Pareto optimality, which is usually labelled as Pareto-Koopmans optimality.

An appropriate benchmark is required to assess the productive efficiency of a DMU in comparison with the performance of other DMU's in the same branch. The original article by Farrell (1957) introduced the concept of the best practice reference frontier. Based on actual observations, this frontier specifies for a DMU the maximum quantities of outputs it can produce given any level of inputs and, for any levels of outputs, the minimum quantities of inputs it needs for producing. Consequently, efficient observations will be positioned on the best practice frontier. Farrell also has developed an indicator of efficiency that measures the distance of inefficient observations from the best practice frontier. For example, in the case of input efficiency one searches for the maximum scalar wise reduction of all inputs yielding the same output. In terms of isoquant analysis, input efficiency is measured along a ray through the origin. Farrell's efficiency measure has been extensively applied in economic literature (see e.g. Färe, Grosskopf and Lovell (1985)).

To determine the best practice frontier, we use two non-parametric methods. The first one is Data Envelopment Analysis (DEA), which originates in the works of Charnes, Cooper and Rhodes ((1978), (1981)). The efficiency of each decision making unit is obtained as the maximum of a ratio of weighted outputs to weighted inputs subject to the constraint that the similar ratios for every DMU be less than or equal to one. This amounts to the following linear programming problem, which must be solved for each DMU:

$$Max\ h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \tag{1}$$

subject to:

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \forall j \tag{2}$$

$$u_r > 0 \tag{3}$$

$$v_i > 0 \tag{4}$$

The particular DMU being evaluated is identified by subscript 0; all other are denoted by j (j = 1,...,n). The maximum of the objective function, h0, is the DEA efficiency score assigned to DMU_0. Each DMU uses an m-dimensional input vector, x_{ij} (i = 1,...,m), to produce an s-dimensional output vector, y_{rj} (r = 1,…,s). u_r and v_j are the output and input weights respectively and constitute the variables of the problem.

This non-linear problem can be transformed to a tractable linear programming problem (Charnes, Cooper and Rhodes (1978)). For computational purposes, this model may be replaced with its equivalent dual formulation:

$$Min \quad h_0 = \tau_0 - \varepsilon \sum_{i=1}^{m} S_i^- - \varepsilon \sum_{r=1}^{s} S_r^+ \tag{5}$$

subject to

$$\tau_0 x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - S_i^- = 0 \tag{6}$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j - S_r^+ = y_{r0} \tag{7}$$

$$\lambda_j, S_i^-, S_r^+ \geq 0 \tag{8}$$

An arbitrarily small positive number, introduced to ensure that all of the observed inputs and outputs have positive values and that the optimal value h0 is not affected by the values assigned to the so-called "slack-variables" $(S_i^- + S_r^+)$ and. The latter variables are associated with the input and output inequalities. If they have any positive components then it is possible to increase the outputs or reduce the inputs without violating any constraint.

DEA comprises several models depending on the assumptions which are made about the nature of the returns to scale. The variable returns to scale condition (DEA-V) occurs if $\Sigma\lambda = 1$. No restriction on $\Sigma\lambda$ corresponds to allowing only constant returns to scale.

The second approach was developed by Deprins, Simar and Tulkens (1984) and is referred to as the Free Disposal Hull (FDH) method. FDH only rests on the assumptions of strong disposability of outputs and free disposability of inputs. The latter assumption rules out that an increase in inputs results in a decrease in outputs. Strong disposability of outputs implies that any reduction in outputs remains producible with the same amount of inputs.

To calculate the individual efficiency measure of each DMU a data classification algorithm based on vector dominance reasoning can be used. It proceeds as follows. Each observation is sequentially compared to all others. A DMU is declared inefficient if it is possible to find another observation, which produces the same, or more outputs with strictly less of at least one input, or which uses the same or less inputs to produce strictly more of at least one output. In this sense they are dominated by at least one other observation. Input-output combinations which are undominated, are declared efficient. DMU's which are efficient but which never dominate another observation are called "efficient by default". More formally, to calculate individual efficiency measures the linear programming problem mentioned before can be used, adding the condition that $\lambda_j \in \{0,1\}$.

Figure 1 illustrates the FDH and DEA-V best practice frontier and the radial efficiency measure for the two inputs-one output case. The full dots represent actual input combinations which produce the same output. The solid and dotted line represent the DEA-V and FDH best practice frontier respectively.

In the FDH setting the three DMU's a, b and c are declared efficient and shape the stepwise best practice frontier. These three DMU's are not dominated by any other observation. Moreover, DMU a is declared efficient by default as a does not dominate any other DMU. DMU b dominates four other observations and DMU c dominates five other DMU's. DMU k is dominated both by b and c, but the best practice reference is identified as c. This follows from the radial distance measure which calculates the input efficiency score of DMU k by the ratio ok"/ok.

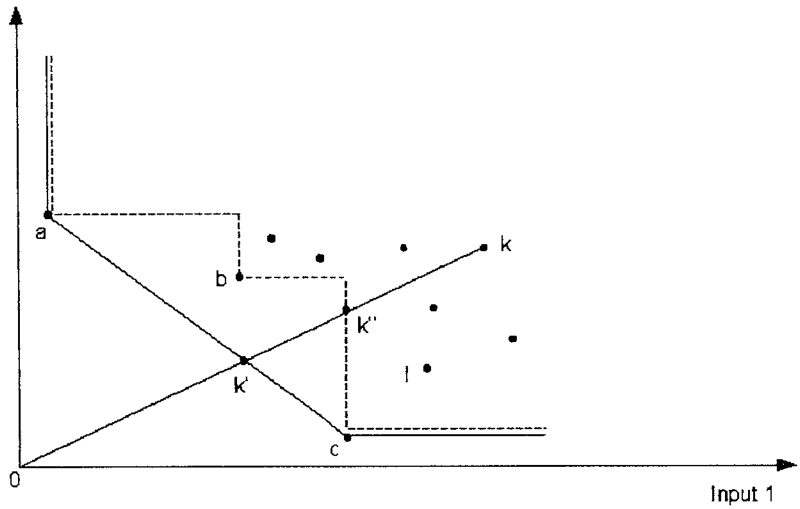
Within the framework of DEA-V the best practice frontier is represented by the solid line. In this reasoning only two DMU's, i.e. a and c, are considered to be technically efficient. For observation k the input efficiency will be lower as it is now measured by the ratio ok'/ok.

Generally speaking the FDH method yields the more generous reference frontier as FDH encompasses the actual data as close as possible. Under FDH there will be a larger number of efficient DMU's, but also the individual efficiency scores of the inefficient DMU's will be higher than under DEA-V.

One of the attractive points of the FDH reference frontier is its strong intuitive appeal. Inefficiencies are calculated in comparison with actually observed input-output combinations. From a management point of view this enhances the credibility of this approach when compared with other methods where inefficiencies are calculated with respect to a hypothetical point on the best practice frontier (see De Borger, Kerstens, Moesen and Vanneste (1994)).

FIGURE 1

Input efficiency for FDH and DEA-V



A major consideration is the sensitivity to outliers. It can be argued that this sensitivity increases with the number of assumptions that are imposed on the best practice frontier. Consequently, in comparison with other non-parametric methods, the FDH reference technology will be the least sensitive to the presence of outliers. A second comment relates to the impact of both the number of observations and the number of input and output dimensions considered in the FDH-approach. An increase in the input and output dimensions increases the probability of efficiency, since the probability of being dominated over all dimensions decreases. Increasing the sample size on the other hand, increases the possibility of dominance and therefore the probability of being declared inefficient (Tulkens (1986)).

Finally, discontinuities in the best practice frontier can lead to major differences in calculated measures of inefficiency even though observations have more or less the same input-output structure.

III. DATA

In Belgium, the data concerning the activities of the tax offices are collected by the Ministry of Finance. After the final audit of a tax return, each audit officer registers his activities in detail. The individual reports are aggregated on a monthly basis for each tax office. This results in output data such as the number of audited returns, the number of under-declarations, the number of control visits in loco, fines, ...

Time and effort to audit a tax return vary considerably across categories. Wage-earners typically file a return which is less complex than the tax declaration of independent professionals who itemise their incurred expenses (see e.g. Frompton (1993)). Fortunately, the Belgian statistics report for each tax office separately the more routine-type returns of category A and the more complex returns of category B. To take account of the "quality" of the audit we like to add, for each category, the number of audits that lead to an increase in the tax base.

The main inputs are labour, capital and materials. Wages are the dominant cost component amounting to 80 pct. of the total operating costs. For the present study we could only obtain data on personnel, expressed in full-time equivalents. However, common administration standards make the allocation of office space and equipment (computer terminals, furniture, ...) to a large extent proportional to labour input.

In the remainder of this paper, the following variables are used:

- *Output:* – the number of audited returns of category A;
  – the number of audited returns of category A that lead to an increase in the tax base;
  – the number of audited returns of category B;
  – the number of audited returns of category B that lead to an increase in the tax base.
- *Input:* personnel, expressed in full-time equivalents.

In Belgium the audit of the returns for the personal income tax is carried out by 313 regional tax offices. Incompleteness of the data forces us to restrict the investigation to 289 tax offices. The data refer to the fiscal year 1991, recorded at the 30th of June 1992.

A summary of the data is presented in Table 1. Taken together the 289 tax offices employ 2,815 people, auditing some 354,400 returns of category A and 534,600 returns of category B. For the more complex B-type returns one out of three returns leads to an expansion of the tax base. This is only 12 pct. for the more routine-type A-returns.

On average an audit office has a staff of 9 to 10 full-time employees, treating 1,226 A-returns and 1,850 B-returns. This amounts to a "productivity" per employee of 126 A-returns and 190 B-returns.

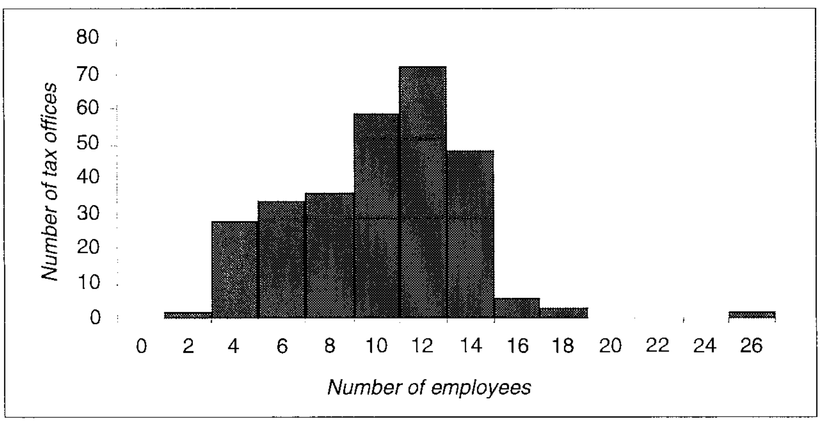
TABLE 1

Descriptive statistics

	Output				Input
	Number of audited return	Increases in the tax base	Number of audited return	Increases in the tax base	
	Category A	Category A	Category B	Category B	Personnel
Total	*354,393.00*	*42,303.00*	*534,603.00*	*177,033.00*	*2,814.50*
Max	10,256.00	1,234.00	354,393.00	42,303.00	66.00
Min	56.00	2.00	670.00	90.00	2.00
Average	1,266.27	146.38	1,849.84	612.57	9074
Standard edivation	964.16	151.07	655.10	281.24	5.35

FIGURE 2

Size distribution of input



The dispersion in input and output size between the regional audit offices is rather substantial. For each variable the standard deviation is reported in Table 1. The distance between the maximum and minimum observation is huge. However for the input size this discrepancy is largely due to 1outlier at the bottom and 2 outliers at the top of the sample distribution. This observation is confirmed in Figure 2 where the size distribution is shown.

TABLE 2
Size distrubution of outputs

Number of audited returns category A		Number of audited returns category B	
Scale	Distribution	Scale	Distribution
0-300	39	0-300	0
300-600	42	300-600	0
600-900	34	600-900	7
900-1.200	39	900-1.200	41
1.200-1.500	44	1.200-1.500	38
1.500-1.800	33	1.500-1800	58
1.800-2.100	19	1.800-2.100	51
2.100-2.400	20	2.100-2.400	50
2.400-2.700	9	2.400-2.700	23
2.700-3.000	5	2.700-3.000	8
3.000-3.300	1	3.000-3.300	7
3.300-3.600	1	3.300-3.600	2
3.600-3.900	1	3.600-3.900	0
3.900-4.200	0	3.900-4.200	1
4.200-4.500	0	4.200-4.500	0
>4.500	2	>4.500	2

The size distribution of outputs is presented in Table 2. It appears that in small size offices the audit of A-type returns is dominant, whereas the relative share of B-type returns increases for the medium and larger audit offices.

IV. RESULTS

For each tax office (DMU) the efficiency score was calculated following the FDH-method and the DEA-method assuming both variable and constant returns to scale. The efficiency measure indicates the relative radial distance from the DMU to the best practice frontier. An index equal to one means that the tax office is, relatively speaking, technically efficient. A measure less than one indicate the degree of inefficiency. An equiproportional increase of the outputs by the inverse of the efficiency score minus one would convert the DMU into an efficient office. For instance DMU 2 has an efficiency score of 0.88. The inverse of 0.88 minus one equals 0.14. This means

TABLE 3

Efficiency scores

	FDH	DEA-V	DEA-F
Number of efficient DMU's	99.00	21.00	10.00
Max	1.00	1.00	1.00
Min	0.29	0.24	0.22
Average	0.85	0.70	0.60
Standard deviation	0.16	0.16	0.16

that DMU 2 should raise its outputs by 14 pct. in order to achieve efficiency.

The FDH frontier constitutes the closest envelope of the actual data. DEA-V takes an intermediate position, whereas DEA-F represents a more "remote" frontier. Consequently FDH will find more DMU's efficient when compared to DEA-V and DEA-F. The results of Table 3 confirm these inferences. It appears that under FDH 99 DMU's are declared efficient, i.e. 34.3 pct. of the sample. The number of efficient DMU's decreases to 21 (i.e. 7.3 pct. of the observations) under DEA-V. Only 10 tax offices remain also efficient under the strictest version of DEA-F. These are the DMU's: 15, 27, 65, 133, 134, 185, 225, 255, 256 and 275. The average efficiency score of 0.85 is rather high for FDH, but decreases to 0.70 and 0.60 for DEA-V and DEA-F respectively. The lowest individual efficiency measure varies between 0.22 for DEA-F and 0.29 for FDH.

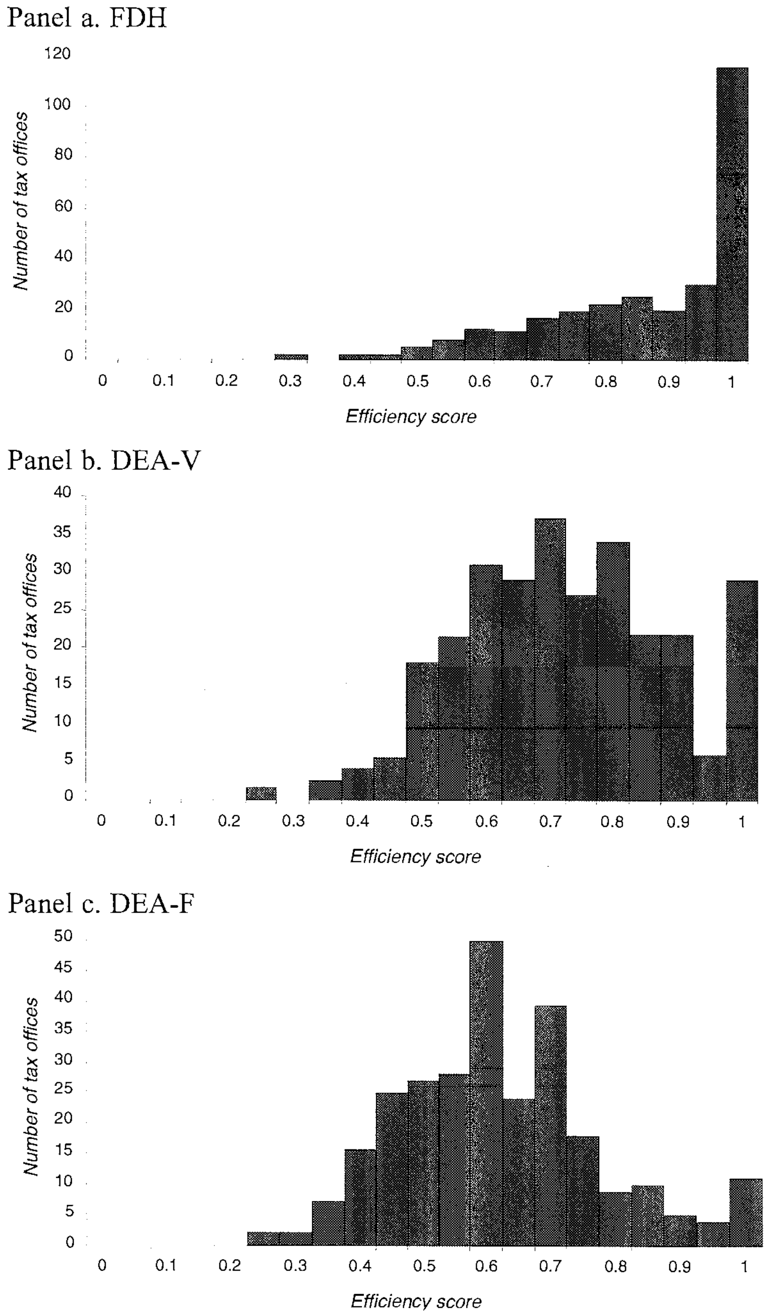
Although the standard deviation of 0.16 happens to be the same for each method, there is a remarkable difference in the distribution of the efficiency scores. Figure 4 represents, for each of the three methods, this frequency distribution. As expected the shape is extremely fat tailed for the more generous FDH-method.

V. SENSITIVITY ANALYSIS

As mentioned before, FDH and DEA are non-parametric methods that use linear programming techniques to determine the best practice frontier. This has the clear advantage that a priori no functional form has to be specified. On the other hand, these deterministic methods are sensitive to outliers in the data.

FIGURE 3

The frequency distribution of efficiency scores



In this section we try to investigate the sensitivity of the DEA-V results to the presence of outliers. Therefore we eliminate the outliers from our main data set of 289 tax offices using one of the procedures outlined in Belsley, Kuh and Welsch (1980). This technique constructs a test statistic based on the so-called leverage value $h_i = x_i(X'X)^{-1}x_i'$ of each observation. The leverage value determines the importance of the observation in the space spanned by all dimensions in the data set. Use of the appropriate test statistic resulted in the detection of 17 outliers. From these outliers 11 were declared efficient in the original analysis.

We repeated the calculation of the DEA-V efficiency scores now based on the data set obtained after deleting the 17 outliers. Dropping the outliers resulted in a marginal increase in the average (from 0.70 to 0.74) and minimum (from 0.24 to 0.27) efficiency measure. The number of technically efficient tax offices remained the same.

The correlation and rank correlation coefficients indicated a value of 0.93, which leads us to conclude that outliers have little impact on the efficiency measures and that, from this point of view, the results can be considered as fairly stable. Consequently we will use the DEA-V efficiency estimates of the original analysis as the left hand variable in the explanatory model of the next section.

VI. EXPLAINING PRODUCTIVE INEFFICIENCIES

The objective of this section is to investigate which variables do influence the efficiency level of the tax offices. An appropriate model to explain efficiency differences should take account of the characteristics of the distribution of these efficiency measures. Efficiency scores larger than one are not observed, meaning that the dependent variable is right censored. Consequently, ordinary least squares will result in inconsistent coefficient estimates. A model that accommodates the specific characteristics of the distribution of the efficiency measures is the Tobit censored regression model, which can be defined in the following way:

$$Y^* = \beta' X + \mu, i = 1, \ldots, n \tag{9}$$

$$Y = Y^* \text{ } if \text{ } Y^* < 1 \tag{10}$$

$$Y = 1 \quad if \qquad Y^* \geq 1 \tag{11}$$

where u is assumed to be normally distributed. The latent variable Y* is not directly observable. Its observed counterpart is the efficiency index Y, which is censored at the limit level of 1, thus making the true value of Y*. For Y* less than one both Y and X are observed while for $Y^* \geq 1$, X is observed and Y equals the limit value of 1.

When the efficiency score is less than one, $t = (Y - \beta'X) / \sigma = \mu / \sigma$ has the standard normal distribution with density function

$$f(t) = \frac{1}{2\pi}\left(-\frac{t^2}{2}\right) \tag{12}$$

and cumulative distribution function

$$F(Z) = \int_{-\infty}^{z} f(t)dt \tag{13}$$

For efficient DMU's it is known that $Y^* \geq 1$ or $\beta'X + 1$. This term is equal to $\mu/\sigma \geq (1 - \beta'X)/\sigma$ with probability function

$$F\left(\frac{\beta'X}{\sigma} - \frac{1}{\sigma}\right) \tag{14}$$

Consequently, the likelihood function of the Tobit model is as follows:

$$LH = \prod_{y_i<1} \frac{1}{\sigma} f\left(\frac{Y_i - \beta'X_i}{\sigma}\right) \prod_{y_i \geq 1} F\left(\frac{\beta'X_i - 1}{\sigma}\right) \tag{15}$$

Maximising this likelihood function with respect to and will result in the required parameter estimates.

One would expect that managerial and organizational characteristics do have an effect on the efficiency degree of tax offices. In fact some tax offices are managed by a high ranking public servant who has the required qualifications. But in the last decade it has been difficult to attract and keep qualified personnel as the prospects in the business sector were more rewarding. This has led to a shortage of high ranking administrators. As a consequence some tax offices are managed by less experienced and qualified civil servants. The management position is captured by a dummy TITU which has the value one if the person in charge has the required qualifications. The internal monitoring of a tax office may also be reflected in the number of fines (FINE) and the number of official assessments (OFFIC), which are associated with its "eagerness". The number of control visits in loco (LOCO) is somewhat ambiguous since it represents a trade-off between quality and quantity of the audit. Nevertheless a positive sign is expected, as for all the other managerial variables, which somehow refer to more dynamic management of the daily operations.

There are also differences at the organizational level. Some tax offices benefit from the services of a Central Tax Office (CTO), while other do not have the same facilities. A CTO provides the automatic handling of several aspects of a tax file, which leaves room for more and better audits. A dummy variable CTO is introduced to capture the existence of a Central Tax Office. The 313 tax offices are regrouped into thirteen geographical directorates. Some directorates are reputed to have a better management culture and cohesion. Other directorates suffer from a higher turnover of personnel. This is e.g. the case in the Brussels area where most of the civil servants get their training before they are appointed as an audit officer in the province. A dummy variable is allocated to twelve directorates (DIR1, ..., DIR12).

The estimated coefficients are reported in Table 4. Taking a significance level of 10 pct., the critical value for the chi-square is 2.7. It appears that the influence of organizational characteristics can not be neglected. The availability of a Central Tax Office (CTO) increases the efficiency score as expected. Six out of thirteen directorates seem to affect the performance of their tax offices. It helps to belong to directorate number 6 (DIR6). On the other hand, one observes a lower performance in the directorates number 1,3,4,8 and 9.The coefficients of the other directorates (DIR2, DIR5, DIR7, DIR10, DIR11 and DIR12) are not significantly different from zero.

Let us now examine the impact of the managerial variables. The efficiency score is highly related to the position of a qualified manager (TITU). Also the number of fines (FINE) is positively related to the degree of efficiency. However, the number of official assessments (OFFIC) is hardly significant. The zeal of a tax office, as reflected in the number of control visits in loco (LOCO), has also a positive influence.

Finally we also wanted to investigate whether the efficiency score is systematically related to the scale of the tax office. As a scale variable we selected the number of people liable for the personal income tax per tax office (LIATAX). The impact seems negligible which is in sharp contrast with the conclusion of another empirical study on the

TABLE 4

Determinants of technical efficiency: Tobit results

Variable	Coefficient	Standard error	Chi-square statistic
INTERCEPT	52.92	5.37	97.11
TITU	7.54	2.62	8.28
CTO	9.82	3.67	7.17
DIR1	−7.65	4.12	3.45
DIR2	−2.54	3.64	0.49
DIR3	−17.19	4.64	13.72
DIR4	−21.64	4.15	27.24
DIR5	4.87	3.55	1.88
DIR6	6.47	3.65	3.14
DIR7	−5.34	4.63	1.33
DIR8	−14.29	3.99	12.85
DIR9	−7.06	3.62	3.80
DIR10	6.08	4.55	179
DIR11	−9.00	5.55	263
DIR12	0.42	4.20	0.009
FINE	0.019	0.011	3.01
OFFIC	0.043	0.027	2.59
LOCO	0.053	0.010	27.11
LIATAX	0.00002	0.0002	0.012

Log Likelihood for normal distribution: −1088.97

performance of tax offices responsible for the business income tax in Belgium (Amez and Moesen (1994)). There it was found that the efficiency estimates are positively related to the scale of the tax office. It turns out that in the larger tax offices it is feasible that the auditors specialize in the control of certain branches such as banking or the petro-chemical industry.

VII. CONCLUSION

In their agenda for research on determining organizational effectiveness Lewin and Minton identify Data Envelopment Analysis as "... a promising mathematic ... to be potentially useful for relating organizational designs to organizational effectiveness" (Lewin and

Minton (1986), p. 532). In the public sector it is tempting to relate the operational performance, either measured by DEA or FDH, to organizational structures such as government versus private ownership, different forms of contracting out or the regulatory environment (for an overview see e.g. Pestieau and Tulkens (1993)).

In this study we focus on the non-parametric measurement of the productive efficiency of tax offices. We mimic an "engineering" approach to tax auditing where physical inputs are converted into non-monetary outputs; i.e. the number of tax audits and the number of corrections for under-declarations for the two major categories of tax files. At the side of the output indicators we have omitted the tax proceeds to mitigate the influence of environmental differences. In fact, variations in regional relative wealth of the taxpayers could then bias the estimates.

It still appears that some tax offices perform substantially better than others. Our investigation reveals that organizational designs do matter such as the presence of a Central Tax Office and to a lesser degree the monitoring system within each Regional Directorate General. Of equal importance is the positive impact of managerial skills. Offices which are daily managed by a qualified (higher ranking) civil servant seem to perform better on average. Which is an argument for investment in human capital also in the public sector.

REFERENCES

Amez, L. and Moesen, W., 1994, De taxatiediensten van de vennootschapsbelasting, *mimeo,* (Center for Economic Studies, K.U.Leuven).

Belsley, D., Kuh, E. and Welsch, R., 1980, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, (John Wiley, New York).

Bjurek, H., Hjalmarsson, L. and Försund, F.R., 1990, Deterministic Parametric and Non-Parametric Estimation of Efficiency in Service Production: a Comparison, *Journal of Econometrics* 46, 213-227.

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

Charnes, A., Cooper, W.W. and Rhodes, E., 1981, Evaluating Program and Managerial Efficiency: an Application of Data Envelopment Analysis to Program Follow Through, *Management Science* 6, 668-697.

De Borger, B., Kersten, K., Moesen, W. and Vanneste, J., 1994, Explaining Differences in Productive Efficiency: an Application to Belgian Municapilities, *Public Choice* 2, 339-358.

Deprins, D., Simar, L. and Tulkens, H., 1984, Measuring Labor Efficiency in Post Offices, in: Marchand, M., Pestieau, P. and H. Tulkens, eds., The Performance of Public Enterprises: Concepts and Measurement, (Elsevier, North Holland, Amsterdam).

Färe, R., Grosskopf, S. and Lovell, C.A.K., 1985, The Measurement of Efficiency of Production, (Kluwer-Nijhoff Publishing, Boston, Dordrecht, Lancaster).

Farrell, N.J., 1957, The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society* A 120, 253-290.
Frompton, D., 1993, Practical Tax Administration, *Fiscal Publications*, (Bath).
Gulledge, T.R. and Lovell, C.A.K., eds., 1992, International Applications of Productivity and Efficiency Analysis, (Kluwer Academic Publishers, Boston, Dordrecht, Londen).
Koopmans, T.C., 1951, An Analysis of Production as an Efficient Combination of Activities, in: Koopmans, T.C., ed., Activity analysis production and allocation, Cowles Commission for research in economics, Monograph no. 13, (John Wiley and Sons, New York).
Lewin, A.Y. and Minton, J.W., 1986, Determining Organizational Effectiveness: Another Look, and an Agenda for Research, *Management Science* 5, 514-538.
Pestieau, P. and Tulkens, H., 1993, Assessing and Explaining the Performance of Public Sector Activities: Some Recent Evidence from the Productive Efficiency Point of View, *Finanz Archiv* 3, 293-323.
Sandford, C., Godwin, M. and Hardwick, P., 1989, Administrative and Compliance Costs of Taxation, *Fiscal Publications,* (Bath).
Slemrod, J., 1992, Why People Pay Taxes: Tax Compliance and Enforcement, (Press Ann Arbor, University of Michigan).
Smith, A., 1976, Inquiry into the Nature and Causes of the Wealth of Nations, (Clarendon Press, Oxford).
Thiry, B. and Tulkens, H., eds., 1988, La performance économique des sociétés Belges de transports urbains, (CIRIEC, Charleroi).
Tulkens, H., 1986, La performance productive d'un service public: définitions, méthodes de mesure, et application à la Régie des Postes de Belgiques, *L'Actualité Economique, Revue d'Analyse Economique* 62, 2, 306-335.
Tulkens, H., Thiry, B. and Palm, A., 1988, Mesure de l'efficacité productive: méthodologies et applications aux Sociétés de Transports Intercommunaux de Liège, Charleroi et Verviers, Chapter 2 in Thiry and Tulkens (1988).