

Olga Gajda-Lüpke¹

DG for Finance, Budget Unit, European Parliament, Luxembourg

Performance measurement methods in the public sector

Abstract: The notion of performance measurement (PM) in the public sector emerged in the framework of the New Public Management paradigm and public management reforms in industrialised as well as in developing countries. The paper discusses the following issues regarding PM – the definition of performance and types of measures, purpose of the PM and possible threats while applying PM. Firstly, as regards the types of measures - it is argued that measures of efficiency and effectiveness should be preferred over output measures. Secondly, the purposes of PM are: informing users in the budget cycle, better planning, improved management, increased transparency and last but not least PM can motivate future performance. Thirdly, concerning the threats to the performance measurement: gaming, imprecise objectives and unprepared transfer of the new management system can make PM unreliable. The purpose of this paper is to overview and assess the recent developments concerning performance measurement and its methodology. It is also meant to add to the discussion on performance budgeting at the moment of the evaluation of Poland's first performance budget in 2008.

Keywords: performance measurement, performance measures, New Public Management, public sector, effectiveness, efficiency, gaming.

JEL codes: H11, H50, L32.

1. Introduction

The notion of performance measurement emerged in the framework of the New Public Management paradigm and public management reforms. New Public Management and performance measurement have been applied both in industrialised countries and in developing countries including India, Jamaica and Thailand. Recently, it was attempted in Japan and Latin American countries. Also one of the

¹ The author is a Ph.D. candidate and would like to thank anonymous referees for comments and Didier Klethi, Head of the Budget Unit in the European Parliament for his encouragement and support.

key objectives of the EU reform in 2000 was to focus its management on obtaining results on its policy priorities. Poland has introduced the performance (results-oriented) budget in 2008 for part of its expenditure and foresees that this method will extend to the whole expenditure in 2011.

Comprehensive research is necessary to adjust the models of performance measurement to the specific conditions of different countries. A reform including New Public Management approach or only performance measurement should be preceded by a profound analysis of the internal and external environment of the receiving country compared to the country where the model comes from.

The purpose of this paper is to overview and assess the recent developments in performance measurement and its methodology. It is also meant to add to the discussion on performance budgeting at the moment of the evaluation of Poland's first performance budget in 2008. The overview and assessment of the related literature is presented, in order to capture its developments relevant for the practice of the public sector. The first section defines the performance measurement and its different approaches. The second one focuses on various types of performance measures. The third section summarises how the performance information may be used in the decision-making process. Section 4 of the paper focuses on possible problems in using the performance measurement.

2. Defining performance

The notion of performance measurement (PM) emerged in the framework of the New Public Management (NPM) paradigm and public management reforms. Economic theories of organisation and government (New Institutional Economics) constituted the intellectual foundations of the New Public Management, Barzelay (2001, p. 160). According to Hood (1991, p. 4) after Yamamoto (2003, p. 4), NPM encompasses the following seven doctrines:

- 1) an emphasis on *hands-on professional management skills* for active, visible, discretionary control of organisations (freedom to manage);
- 2) *explicit standards and measures of performance* through clarification of goals, targets, and indicators of success;
- 3) a shift from the use of input controls and bureaucratic procedures to rules relying on *output controls* measured by quantitative performance indicators;
- 4) a shift from unified management systems to *disaggregation* or decentralisation units in the public sector;
- 5) an introduction of greater *competition* in the public sector so as to lower costs and achievement of higher standards through term contracts, etc.;

- 6) a stress on *private-sector-style management practices*, such as the use of short-term labour contracts, the development of corporate plans, performance agreements and mission statements;
- 7) a stress on cost-cutting, efficiency, *parsimony* in resource use, and “doing more with less”.

The doctrines (2), (3), (6) and (7) refer explicitly to the concept of performance by mentioning indicators, targets and putting emphasis on outputs. NPM calls for a precise definition of goals, targets, and indicators of performance measurement, Yamamoto (2003, p. 5). According to Pollitt and Bouckaert (2004, p. 8), the recent public management reforms focus on deliberate changes in the structures and processes of public sector organisations in order to improve their functioning. These reforms can serve different purposes, among others: savings in public expenditure, improving quality of public services, making the operations of the government more efficient and increasing the effectiveness of the policies to be implemented, Pollitt and Bouckaert (2004, p. 6). Thus performance measurement defines the methodology to quantify efficiency and effectiveness and gives to management a tool to improve functioning of an organisation. Measuring performance in the public sector is not new. The UK and the USA measured performance of teachers already in the nineteenth century, Pollitt and Bouckaert (2004, p. 90). Performance indicators for government activity can be traced back to the Napoleonic times and further back to the invention of writing, accounting, and statistics, Caiden and Caiden (2006, p. 142). Currently the interest in this field is increasing and the notion is becoming more complex as more levels, fields and management functions are included. Not any more for internal use only, it is required by members of legislative bodies and becomes accessible to the public. Nowadays PM concerns mostly decision-making, controlling and providing accountability, Bouckaert (1996b, p. 234) after Pollitt and Bouckaert (2004, p. 90). In addition, it can be used for service improvement, Ammons and Rivenbark (2008, p. 308).

Performance measurement has been defined by OECD, Curristine (2005, p. 131) as evidence for performance-related results such as economy, efficiency, and effectiveness. It may include output and outcome targets and goals, and key processes that affect them. In addition, the evidence may show how the outcomes are attributed to a specific programme or government activity. Performance measures or performance indicators can be quantitative or qualitative and are normally compared against some base or standard.

When considering the product of a public service, taxpayers, users and citizens have usually diverse interests. For the taxpayers, it is the cost of the service that matters, whereas the users of public service are interested rather in the quality of the service. The citizens, on the other hand, focus on the impact of the service on their general well-being. The concept of performance has been developed to reconcile these various perspectives. It examines public services from the point of view of “in-

puts”, “outputs”, “outcomes” and “impact”. Thus, public administrators are required to apply performance indicators not only in terms of appropriations used or human resources deployed but also in terms of the added value for society as a whole, ECA (2008, p. 5). In sum, the notion of performance was present for a long time in the practice of public administration. It found its significant place in the theory of the New Public Management and continues to develop reconciling perspectives of taxpayers, users of public service and citizens. In short, performance measurement is an evidence for performance-related results.

3. Types of performance measurement

Level of performance depends on the measures chosen. Managers are often faced with redundant measures of the same output, each of which may be preferred by a different political principal or stakeholder. Moreover, the managers’ choice of measures can have serious consequences for the estimation of agency problems and the success of programmatic solutions, Nicholson-Crotty et al. (2006, p. 101). The ambiguity of goals in public organisations, along with nature of performance measurement, means that multiple measures or methods for measuring the same concept may be identified as appropriate or correct, Nicholson-Crotty et al. (2006, p. 102). Following OECD and Ammons and Rivenbark (2008), in this paper the performance measures will be divided into the following categories:

- output measures,
- higher-order measures,
- evaluations.

These categories are further discussed below. Furthermore, a comparison of raw versus adjusted measures is examined.

3.1. Output measures

The output (workload) measures are the most elementary measures. Outputs are “goods or services which government organisations provide for citizens, businesses and/or other government organisations/bodies”, Curristine (2005, p. 130). These measures record only the number of units of service produced in a given period of time. They answer the easiest question: How many? They are, however, ill suited for answering more managerially challenging questions: How efficiently? How effectively? Of what quality? A natural advantage of output measures is that their collecting imposes minimal disruption and expense on operating departments. Nevertheless, it has been questioned whether investment in performance measure-

ment, restricted entirely to workload measures, produced any operating benefits at all, Ammons and Rivenbark (2008, p. 307).

These indicators support accountability and, in particular, performance reporting, but are unlikely to yield much of a service improvement. Such a system produces feedback having very little managerial or policy value to operating officials or government executives. It merely documents whether demand for a service is up, down, or relatively stable (e.g. that the library had 100 visitors per day), Ammons and Rivenbark (2008, p. 308). As described in Curristine (2005, p. 130), an output target may be set for these indicators.

3.2. Higher-order measures

Higher-order measures are more sophisticated than output measures. They are also called measures of efficiency and effectiveness, Ammons and Rivenbark (2008, p. 308) or measures of economy, efficiency and effectiveness, Curristine (2005, p. 131).

3.2.1. Definition of economy, effectiveness and efficiency

As defined by the EU (2002, p. 12), the principle of economy requires that the resources used by the institution for the pursuit of its activities are made available in due time, in appropriate quantity and quality and at the best price. “The principle of efficiency is concerned with the best relationship between resources employed and results achieved. The principle of effectiveness is concerned with attaining the specific objectives set and achieving the intended results”, EU (2002, p. 12). The three “Es” constitute the principle of sound financial management in the European Union’s Financial Regulation.

As far as efficiency is concerned, two dimensions should be considered: 1) technical or operational efficiency and 2) allocative efficiency. The technical or operational efficiency refers to the output-input ratio compared to an optimal ratio (hence always less than 100%)². Here output- and input-oriented efficiency can be identified. Output efficiency will aim at maximising output for a given set of inputs, whereas input efficiency will focus on minimising the use of inputs for a given set of outputs. Allocative efficiency refers to the use of inputs in optimal proportions given their respective prices and production technology. Economic efficiency constitutes the product of technical and allocative efficiency, Van Dooren (2008, p. 4). As such the allocative efficiency reflects the principle of economy defined above. Efficiency and productivity are sometimes used interchangeably.

² On the other hand, Ketelaar, Manning et al. (2007) after World Bank (2008, p. 6) define efficiency as a ratio of costs to output, whereas productivity is defined as a relation of output to input.

The measure of output efficiency is applied where there are preset output levels. For instance, citizens are entitled to a number of public services (hospital, unemployment agency, court). Here only inputs can be minimised, Van Dooren et al. (2008, p. 4). Nonetheless, a high quality of service should be attempted. The minimisation of inputs could take place only under the condition that a desired effect is achieved (high measure of effectiveness). Van Dooren et al. (2008, p. 4) takes as an example of input-efficiency with a preset input level, an authorised budget in cultural institutions or environmental agencies. Such a budget is an expression of a political desire that a specific amount of resources is used to reach some social outcome.

3.2.2. Efficiency measures

Efficiency measures are one type of the higher-order measures. They should report with precision the relationship between production outputs and the resources consumed to produce these outputs, as defined above. Measurement of efficiency requires quantitative information on both input and output sides of public service provision, Van Dooren et al. (2008, p. 4).

Measuring output captures the quantity and quality of products and services, where quality is understood as the degree to which outputs contribute to outcomes. We can distinguish between micro-level measures (for individual organisations), intermediate measures (at a sectoral level) and macro-level measures at a government level, Van Dooren et al. (2008, p. 4). A method for aggregation poses some problems since there are no obvious weights of individual outputs for the final government output. On the other hand, there are intangibles like, for instance, process based activities in a diplomatic service, such as negotiating or interest representation. There exists a risk of focusing on the measurable outputs and neglecting the other ones, Shepard (1990), Smith and Rothenberg (1991) after Van Dooren et al. (2008, p. 5). Such gaming effects are discussed further in section 4 of this paper.

Measuring inputs seems more straightforward but it can still be a challenge. Accounting systems can be regarded as institutionalised measurement systems for the input side. It is argued that accruals-based accounting should be used instead of cash-based accounting to reflect better the economic reality of the organisation, Van Dooren et al. (2008, p. 5-6). The cash-based accounting and budgeting is a method of bookkeeping that recognises a transaction when cash is received or when cash is paid. The accruals-based accounting recognises a transaction when the activity (decision) generating revenue or consuming resources takes place, regardless of the fact, when the associated cash is received or paid, Blöndal, (2004, p. 104). To assure comparability of efficiency measures, the cost-accounting systems should follow similar methodology in various public sector organisations. Only then a possible potential of improvement could be identified. One step to guarantee comparability of cost-accounting methods in public sector is the development of International Public Sector Accounting Standards. They are designed

to apply to the general purpose financial statements of all public sector entities, IFAC (2008, p. 17).

Some local governments desiring efficiency measures cope with inadequate cost-accounting system by using another denominator than currency. Instead, they apply staff hours, labour hours or full-time equivalent (FTE) positions³. For example, €10 per application processed, 20 applications processed per €100, 2 applications processed per staff hour. Alternatively, Ammons and Rivenbark (2008, p. 309) suggest applying measures, depicting the extent to which equipment, facilities, and personnel are fully utilised, as well as measures that estimate only roughly the efficiency of production processes (e.g. percentage completed on schedule).

However, even this measure of efficiency becomes complex, when the time of a given employee has to be divided among his different tasks and different outputs. Hence, many local governments resort to reporting “FTE per 1 000 of population or “cost per capita” for services overall or for services of a particular department. It is doubtful though, if these measures could be counted as efficiency measures at all. As Ammons and Rivenbark (2008, p. 310) emphasise, a city government that has no responsibility for fire services because these are handled by a county government will appear more efficient in a total cost per capita comparison than its full-service counterparts that have responsibility for these functions. Per capita cost comparisons on a function-by-function basis reduce this problem, however, they may be imprecise due to different cost-accounting used. Hence, cities unsure of uniformity in cost-accounting systems to their counterparts are recommended to rather focus on measures of quality and effectiveness.

3.2.3. *Measures of effectiveness and outcome measures*

The measures of effectiveness and outcome measures define the extent to which objectives are achieved. They can be measured at different levels of objective setting (e.g. government level, ministries level, units). Outcomes are defined as “effects and consequences of government actions on a community”. They can reflect the intended and unintended results of government actions. For example, an outcome for Ministry for Foreign Trade could be signing of a bilateral trade agreement. Similarly as for output measures, here also targets may be set, Curristine (2005, p. 130).

The measures of efficiency and effectiveness inspire managers, supervisors, and frontline employees to diagnose the problem, if one exists, and to devise strategies to correct it. Thus, they inspire managerial thinking, Ammons and Rivenbark (2008, p. 308). An example of such measures, described in KPRM (2007, p. 69), is a target of the Polish Ministry of Justice to decrease by 20% the amount of court cases handled over twelve months.

³ According to OECD (2008b), “full-time equivalent employment is the number of full-time equivalent jobs, defined as total hours worked divided by average annual hours worked in full-time jobs”.

3.3. Raw versus adjusted performance measures

Performance of a public institution or a government cannot be analysed in isolation. Recently a model has been tested to assess how such factors as: political environment, stakeholder participation, organisational support, and training affect the adoption and managerial effectiveness of performance management, Yang and Hsieh (2007, p. 861). The possibility to adjust the performance measures for the environmental factors is present in the literature. Raw performance measures (RPMs), such as test scores or success rates, are often inferior to performance measures adjusted for client and environmental characteristics, or adjusted performance measures. Most RPMs reflect not only the competency and capability of public organisations but also a broad range of factors over which organisations have little control, such as the characteristics of the clients they serve and the environment in which they operate. Obviously, however, care must be taken over interpretation of such measures, Rubenstein et al., (2003, p. 607). Fair comparability of public sector organisations could be assured with the application of adjusted performance measures. Nonetheless, the concept seems still rather theoretical.

3.4. Evaluations

Evaluation refers to the assessment of an ongoing or completed project, programme, activity or policy. Evaluations of a programme or policy can concern its objectives (policy priorities), design, implementation (process), and results (output/outcome/cost effectiveness). The aim can be to determine the relevance and fulfilment of objectives, in terms of efficiency, effectiveness, economy and sustainability. Various types of evaluations include: ex post/ex ante evaluations, desk evaluations, impact or in-depth evaluations, and self-assessments, Curristine (2005, p. 129).

Evaluation and performance measurement are different tools. While performance measures concern outcomes, outputs and/or process indicators, evaluation often includes a more detailed review of attributes and causality issues. Thus it concerns procedures and processes and typically includes recommendations on changes in activities or programmes to improve performance, Curristine (2005, p. 129), being a complementary tool to the measures discussed in sections 2.1 and 2.2.

In this section a classification of performance measures into output and higher-order measures has been suggested. An alternative classification has been proposed in Lubińska ed. (2007, p. 329). There output measures, outcome measures and impact measures are distinguished. This categorisation relates to the budget classification chosen for the first Polish performance budget, established in 2008. The expenditure has been divided into sections, tasks and subtasks according to COFOG (Classification of Functions of Government), Lubińska ed. (2007, p. 323). Tasks and subtasks are responsibilities of relevant units of administration and reflect their ac-

tivity, Lubińska et al. (2008, p.3, to be published). The impact, outcome and output measures are defined as follows, Lubińska ed. (2007, p. 329-330):

- Impact measures analyse long-term consequences of task implementation. They may measure direct results of task implementation, but also the impact of this task on groups other than the target group. They are defined only as measures of effectiveness. Example: unemployment rate, the number of committed offences.
- Outcomes measures assess the result of activities at the subtask level and measure direct consequences of activities and analyse them from the efficiency and/or effectiveness point of view. Outcome efficiency measures show efficiency of the task implementation, for example lower treatment cost resulting from pro-health activities. Outcome effectiveness measures evaluate fulfilment of task objectives, for example reduced driving time as result of road repairs.
- Output measures reflect the performance of a task or subtask in a short term and show specific public goods or services. They concern direct consequence of activities. Output measures concern both efficiency and effectiveness. Output efficiency measures reflect task performance or task implementation progress from the financial point of view, for example the cost of building one kilometre of a highway. Output effectiveness measures are physical measures that rely on the measurement of the quantity or quality of produced goods or services and show task implementation progress and the technical quality or the process quality. For example the completed number of transplantations compared to those planned.

4. Purpose of performance measures

The size of the public sector and its role in the economy is a major reason why performance of governments should be measured. The impact of government emerges in the quality and nature of goods and services it provides, its redistributive activities, and in the nature of its regulation of market and individual behaviour, OECD (2007b, p. 2). General government expenditure amounts to around 46% of Gross Domestic Product for EU-27 countries in 2007, Eurostat (2008).

The purpose of some recent management reforms in the framework of New Public Management might be set equal to the purpose of the performance measurement, since the performance measurement represents the core of such reforms, as discussed in section 1 of this paper. Pollitt and Bouckaert (2004, p. 6) define the following aims of the reforms: savings in public expenditure, improving quality of public services, making the operations of the government more efficient and increasing the effectiveness of the policies to be implemented. These objectives can be reached through the use of performance information in the decision-making chiefly about financial

and human resources. Using performance measures is more than simply reporting measures or somewhat vaguely considering measures when monitoring operations, Ammons and Rivenbark (2008, p. 305). It can be used to keep or eliminate programs based on how efficient or effective they are. Actual use of performance measures requires evidence of an impact on decisions at some level of the organisation.

Further in this section the following objectives of performance measurement will be discussed:

- informing users in the budget cycle,
- better planning,
- improved management,
- increased transparency,
- motivating performance.

4.1. Informing users in the budget cycle

The budget cycle can be defined as consisting of the following stages: drawing up, implementation, evaluation, audit and discharge of the budget. During each phase of the cycle, performance information (PI) informs different users. In some Latin American countries they are: Presidency, Planning Agency, Ministry of Finance, Line ministers, Congress, External Audit and the Public. In the drawing up phase, PI can be applied in the negotiations between the Planning Agency, Ministry of Finance and Line Ministries. Demonstrable results are also a key input during negotiations to secure budget approval from Congress at the level of the Presidency, World Bank (2008, p. 14). Moreover, in the implementation phase where there are changes to the authorised budget, performance information can be demanded by the authorising body, World Bank (2008, p. 14). As far as audit and discharge phase are concerned, some external audit institutions have also strived to include performance audit in their audit scope on the top of financial and compliance issues, World Bank (2008, p. 17). Also in the European Union, the European Court of Auditors (ECA) conducts a performance audit, verifying whether the financial management of the European Union has been sound, that is if it complies with the principles of economy, efficiency and effectiveness (as defined in section 2.2.1 of this paper). The European Parliament uses the reports of the ECA as a reference for the discharge procedure over the EU budget.

Performance budgeting (PB) and management initiatives seek to shift the focus of decision making away from inputs towards results and to provide mechanisms to improve efficiency as well as to monitor the performance of ministries and agencies (OECD (2007a, p. 24). The efficiency gain is achieved here through better planning of financial resources based on the performance information gathered. According to the OECD definition, performance budgeting is a budgeting that relates funds allocated to measurable results, OECD (2007a, p. 20).

Depending on the approach of PB adopted, countries seek to link performance information to decision on resource allocation tightly, loosely, or not at all, OECD (2007a, p. 42). Thus, OECD distinguishes three categories of performance budgeting in accordance with the use of outputs and/or outcomes measures in resource allocation, Curristine (2005, p. 102-103). These approaches are defined below.

- Firstly, in the framework of presentational performance budgeting, performance information is not linked to the decision on resource allocation. Performance measures are presented in budgeting or other government documents and this information is included as background information only. It does not play a role in the decision making on allocations nor is it necessarily intended to do so, Curristine (2005, p. 102-103). Producing performance information is a free decision of spending ministries. PI, is used mainly for accountability purposes and outside the budget negotiation process rather than as part of it, OECD (2007a, p. 42). This type of PB is, for instance, applied in some public sector areas in Denmark or Sweden.
- Secondly, performance-informed budgeting presents a loose link. It relates fund allocation to measurable results in the form of outputs and/or outcomes. Resources are related to results in an indirect manner. In this method, performance information is very important in the decision-making process but it does not necessarily determine the amount of resources allocated. Output or outcome measures are actively and systematically used. They shape budget decisions along with other information on performance or other information pertaining to macro restrictions on fiscal policy and policy priorities. Performance information is important, but it is not absolute and does not have a predefined weight in the decisions. The final weightings depend on the particular policy context, Curristine (2005, p. 102-103), OECD (2007a, p. 42). Performance measures can be used for planning and/or accountability purposes, OECD (2007A, p. 42). This approach is applied in most of the OECD countries⁴.
- Thirdly, direct performance budgeting involves the allocation of resources directly and explicitly to units of performance measures. The amount of resources granted can be based on a formula/contract with specific performance or activity indicators. Hence, funding is directly based on the results achieved. In theory, this approach can be applied with or without formulas. In the case of no formulas, there is a procedure or contract for systematically providing funding rewards or penalties on the basis of performance against targets. Here better performance always means more resources and worse performance always means less, but the amount of appropriation is not defined. Generally, however, the direct linkage approach

⁴ OECD Member countries are: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States, OECD (2008c).

is used with formulas, Curristine (2005, p. 102-103). This approach is present in the sector of education in Chile, Denmark, Finland, Hungary, Iceland, Norway, Portugal or Sweden. It is also applied to the health sector in Chile, Denmark, Iceland, Norway, Portugal or the UK, OECD (2007a, p. 47).

Certain budget classifications are more conducive to the integration of performance information than others (Pollitt, 2001, p. 18) after OECD (2007a, p. 40). For example, programme or outcome and/or output classifications are more open to incorporating performance information than line-item budgets. Nonetheless, changing the budget structure does not necessarily alter the budgetary decision-making process. It can remain on a traditional incremental basis, especially if the budget process itself and the incentives for the actors in this process have not been altered. Most countries that have changed their budget classification have also tried to change their budget processes. A few countries changed their budget structure to focus on outputs and/or outcomes. For example Australia, New Zealand and the United Kingdom changed their structures as part of their initiatives to introduce accruals-based budgeting, OECD (2007a, p. 40-41). The European Commission introduced Activity-Based Budgeting in 2004. The new nomenclature for the 2004 EU budget followed the structure by policy area and activity used by all Commission's services as the reference structure of their annual management plans, EC (2004, p. 6). Currently, the Financial Framework 2007-2013 and consequently the annual budgets follow the structure of 6 headings. Headings 1-4 correspond to the EU policy priorities, heading 5 represents administrative expenditure and heading 6 refers to compensations linked to enlargement, EU (2006, p. 10). As underlined in ECA (2008, p. 23), a lack of activity-based budgeting makes it difficult, if not impossible, to link actual activities and the implementation of the budget.

4.2. Better planning

The improved setting of objectives and a greater emphasis on planning come as a natural consequence of using performance information in the decision-making process and management. The information is taken into account at the closing stage and the institutions are held accountable for the fulfilment of policy objectives. Therefore they will be more inclined to plan in a reasonable way and set "SMART" objectives, where SMART stands for specific, measurable, achievable, relevant and timed, EU (2002, p. 12), EU (2007, p.19).

Setting goals and objectives, planning program activities to accomplish these goals, allocating resources to these programs, become crucial in view of monitoring and evaluating the results to determine if they are making progress in achieving the established goals and objectives and modifying program plans to enhance performance, Hatry et al. (1990), after Behn (2003, p. 587). According to the World

Bank (2008, p. 5), the objective of performance budgeting is a better allocation of resources consistent with political and social goals.

4.3. Improved management

Achieving savings in public expenditure was a major reason to introduce performance measurement into public sector, Pollitt and Bouckaert (2004, p. 6). This reform was to restrict the public debt increase due to economic, social and structural problems, Lubińska ed. (2007, p. 318). In the EU it is of particular importance as a low debt and deficit need to be maintained as part of the Lisbon strategy, Lubińska et al. (2007, p. 709), Lubińska et al. (2008, p. 1).

As regards improved management and improved monitoring of performance, performance information could help to improve allocative efficiency, i.e. efficient allocation of public expenditure in accordance with government priorities, cf. section 2.2.1 of this paper. The increased efficiency is reached when funds are allocated towards high-performing programmes that are achieving government goals. This can only be done if performance information is actually used in the allocation of resources, OECD (2007a, p. 65). Such performance information is also used for monitoring the progress of the work programme tasks and the extent to which the objectives have been realised, ECA (2008, p. 18).

Some argue that the technical (operational) efficiency may also be increased while using direct performance budgeting, Robinson and Brumby (2005) after OECD (2007a, p. 65). Denmark has claimed that the application of what they term the “taximeter model” in higher education and health has created incentives that – combined with the increased financial flexibility for universities and hospitals – generated efficiency gains, OECD (2007a, p. 65). However, with direct performance budgeting, the public administration is particularly vulnerable to dysfunctional and gaming behaviour, as further discussed in section 4 of this paper. Concern has also been raised about the impact of the direct PB on overall aggregate fiscal discipline. In the case of health care in Norway, the introduction of activity-based financing contributed to an increase in health care expenditure (Anderson, Curristine and Merk, 2006) after OECD (2007a, p. 65).

As underlined by Caiden and Caiden (2006, p. 140-141), political and public accountability should not be sacrificed to administrative expediency, though. Managerial efficiency should not take precedence over other governmental priorities. If the government needs to follow certain objective realised currently with a help of inefficient programme, there will be no possibility to eliminate it. “Government organisations are not just instrumental but, in time, become institutional, with their own vested interests, lobbyists, client groups, professional occupations, modes of operation, traditions and objectives”, Caiden and Caiden (2006, p. 140-141). These groups may game with the method of performance measurement.

As a practical example for improved management, in the case-study by Ammons and Rivenbark (2008, 306-307), efficiency measures for residential refuse collection were examined between cities and some measures were found to be out of line. The measures indicated high unit costs and low worker productivity due to underutilisation of labour and equipment. As a consequence, the city concerned discontinued private refuse collection and extended its own operation into that area without adding equipment or labour, Jones (1997) after Ammons and Rivenbark (2008, p. 307). Hence the city increased their efficiency for this task.

4.4. Increased transparency

Showing results to the public has become a priority of public administration. It came as a response to increasing budget tensions and greater demand by citizens for high quality public services, ECA (2008, p. 5), Lubińska ed. (2007, p. 316). The need for transparency is a starting point to evaluate efficiency and effectiveness in the public sector, Lubińska et al. (2007, p. 705). Citizens can be informed, for example, through benchmarking information for schools or hospitals. Nevertheless, such rating does not explain the underlying causes of good or poor performance as it provides only a snapshot in time. Despite the fact that manipulating the results (gaming) may happen, league tables and benchmarking that do provide such explanations can help citizens to choose from among local schools and hospitals. Such information, while not perfect, could provide some guidance with regard to the level of performance and service provision, OECD (2007a, p. 63). On the other hand, performance measurement can improve citizens' trust in government directly through their participation in the evaluation process or indirectly by improving citizens' perception of government performance, Yang and Holzer (2006, p. 114).

The transparency is quite high through publishing performance measures. Nonetheless, providing meaningful information remains a challenge. Most of the OECD countries, as stated in OECD (2007a, p. 52), publish information on the performance of the public sector⁵ but the information is not always easily understandable and does not facilitate comparisons. For instance, Program Assessment Rating Tool (PART) system in the United States publishes the performance of programmes in a rating system that allows comparison. In OECD countries this approach is adopted rather in the area of local service delivery: league tables for hospitals and schools (UK), benchmarking performance of state services (USA, Australia), public availability of internal audits and programme evaluations (Canada).

⁵ The European Commission managing the operational expenditure of the European Union publishes the results in Annual Activity Reports, EC (2007b).

4.5. Motivating performance

Ministries of Finance, as actors co-ordinating the budgetary process can use performance results to motivate agencies to improve performance. These incentives can be divided into three broad categories: 1) funding, 2) flexibility and 3) public recognition, OECD (2007a, p. 48). In order to reward or punish performance, departments can be given more or less management flexibility, i.e. decrease or increase in internal controls, OECD (2007a, p. 51). Another incentive is making results public and hence name and shame underperformers and reward good performance. The name-and-shame approach often seeks to compare and score the performance of one agency against another, or an agency against its historical record, OECD (2007a, p. 52). This can be done through scorecards and/or benchmarking, or via league Tables, as discussed in point 3.4.

There are, however, only few ministries of finance in OECD countries that use performance information to reward or punish agencies. According to the survey by OECD, only 4% use often the performance measures to eliminate programmes, 10% to cut expenditure and 11% to determine pay. As regards using evaluations for action the percentages amount to 11%, 15% and 5% respectively, OECD (2007a, p. 48). The actual decision-making power can rest with the head of government or result from political negotiations within the cabinet or the legislature. For example, in the United States, the President's 2006 budget proposed twenty one programme terminations and nine programme funding reductions. The Congress accepted only seven of the terminations and four of the reductions. Budgetary decision making takes place in a political context, hence proposals to eliminate programmes or reduce their funding can encounter political resistance, especially if these programmes bring benefits to important political groups or are high political priorities, OECD (2007a, p. 49).

Caution is needed with using the performance information for the purpose of elimination or funding reduction. First of all, the causes of poor performance (which could be based on lack of funding) have to be considered. Otherwise, the situation could get worse and badly performing agencies will continue to underperform. Performance in any given period can be influenced by a variety of factors, both internal and external, that may or may not be controlled by an agency. In addition, in some OECD countries the PI may not be of sufficiently high quality to be used in budgetary decision making in this manner, OECD (2007a, p. 49).

The motivation to increase the efficiency and effectiveness may be strengthened if an agency's performance is directly reflected in pay of senior management. In Denmark and the United Kingdom, the chief executive's performance bonus is partially determined by the agency's success achieving its performance targets, (Ketelaar, Manning and Turkisch 2007) after OECD (2007a, p. 54). Furthermore, it seems that public organizations' managers should measure performance because

such measures are helpful in achieving eight specific managerial purposes: “to evaluate, control, budget, motivate, promote, celebrate, learn and improve”, Behn (2003, p. 586). Applying performance measurement supports improved communication within and across branches of government, advances learned discussion about the results of government activities and services, and adds value to budgeting decisions by providing pertinent information about results, as well as costs and activities, Melkers and Wiloughby (2005, p. 188). Theoretically, as discussed in (OECD, 2005b) after OECD (2007a, p. 50), the introduction of PI into budgeting and management processes should be followed by a relaxation of input controls and increased financial or managerial flexibility in the areas of spending and staffing. In practice, in the OECD countries there is no clear trend to relax input controls while introducing formalised performance information.

5. Threats to reliable performance measurement

5.1. Gaming

Dysfunctional and unintended responses, also known as gaming responses have to be taken into consideration when regarding the implementation of performance measures in public organisations. Performance measures may communicate objectives that may not exactly correspond to the organisations true goal and this misalignment may result in inefficient resource allocation, Courty and Marschke (2003, p. 269). Bevan and Hood (2005, p. 8) define gaming as “reactive subversion such as ‘hitting the target but missing the point’ or reducing performance where targets do not apply”. A mechanical approach to punish bad and promote good performance can generate perverse incentives and encourage agencies to manipulate data. Incentives to provide accurate information are influenced by the expectations of how it will be used in decision making, OECD (2007a, p. 49-50). Hospitals may engage in skimping (not providing the full service), dumping (avoiding the high cost of difficult cases) and creaming (over servicing low-cost, “easy” patients), OECD (2007a, p. 65). In the UK, ambulances were parked outside the accident and emergency departments so that patients were not admitted until they could be treated within the four hour waiting target, Hood (2007, p. 33). In Korea, the only country to attempt to automatically link performance to funding on a government-wide scale, programme ratings are to be used to reduce the budgets of ineffective programmes. An automatic 10% budget cut is to be enforced for ineffective programmes. Nonetheless, problems already appeared as the information received from ministries is of poor quality. There have been issues with gaming, goal distortion and the presentation of misleading information. Most OECD countries have not linked automatically per-

formance to results or resources on a systematic government-wide scale, perhaps realising the difficulties, OECD (2007a, p. 49-50).

Bevan and Hood (2005, p. 8) outline three unintended effects as regards setting targets for performance: ratchet effects, threshold effects and output distortions. They recall that the 60-year history of Soviet targets provides evidence for major gaming problems.

- Ratchet effects refer to the tendency for central controllers to base next year's targets on last year's performance. Managers who expect still to be in place for the next target term tend not to exceed targets even if they could easily do so, Bevan and Hood (2005, p. 9). Hence, targets hold back managers' achievements, since high achievements would mean that targets would be set higher for the next period, Hood (2007, p. 34).
- Threshold effects refer to the way that targets can unintentionally destroy incentives to reach above the target level, Hood (2007, p. 34). Those performing below the target level will strive to reach the target, however those doing better than the target will allow their performance to deteriorate to the standard, Bevan and Hood (2005, p. 9).
- Output distortions denote what happens when managers concentrate on incentivised activities at the expense of others, Hood (2007, p. 34). Distortions including neglect of quality were documented in the Soviet regime, Berliner (1988, pp. 283–284) and Nove (1958, pp. 4-9) after Bevan and Hood (2005, p. 9).

It does not mean, however, that one should give up on target setting completely. Targets have the advantage of focusing attention on priorities set by elected politicians. An intelligent application could be assured by making them less predictable (hence harder to game). Moreover, target reporting could be supplemented by open forums for accountability rather than limiting itself to a bureaucratic scrutiny, Hood (2007, p. 24). Gaming is not particular to the public sector, private companies also need to be cautious in using targets when measuring their performance.

5.2. Imprecise objectives

If the objectives are not set in a specific, measurable, achievable and relevant way at the planning stage, there will not be a possibility to refer them to the results. This problem was among others outlined by the European Court of Auditors in their recent report on EU's agencies, ECA (2008, p. 4). The objectives were often vaguely worded and not properly co-ordinated with the Community sectoral policy priorities, ECA (2008, p. 29). Therefore, setting objectives in the planning phase needs to take into account the later measurement of performance versus the objectives.

5.3. Unprepared transfer of policy

A reform including New Public Management approach or only performance measurement should be preceded by a deep analysis of the internal and external environment of the receiving country compared to the country where the model comes from. Economic, social and administrative development needs to be taken into account Lubińska ed. (2007, p. 30-31, 321). According to Pollitt (2004, p. 1), transfers of management technology usually involve significant adaptation or even a transformation of the imported technology. Secondly, they commonly entail the import of more than just the tool or instrument but packages of assumptions, norms, prerequisites and unintended consequences. Thirdly, it is more accurate to regard transfers of management technology rather as part of an on-going process of adaptation and development as they are not a standard solution to the problems, Pollitt (2004, p. 1-8). Pollitt (2004, p. 5) underlines, that when a transfer of a particular management system or technique is contemplated, one needs to consider how far each of the major factors - culture, political system, pattern of institutions, management strategy, nature of the primary task - matches between the 'exporting' and 'importing' jurisdictions. The more mismatches, the more probable it is that the transfer will fail or produce unexpected results.

6. Conclusions

Reforms of public administration in industrialised as well as in developing countries to adopt New Public Management are a general trend. The reforms will differ from one another as there is no standard solution that fits all. Nonetheless, certain milestones seem common. One of them is the attempt to measure the performance of the public sector. Some countries have already a few decades of experience in this field. Poland and other new market economies need to work further on adapting the performance measurement techniques in the public sector considering the solutions from different countries and adjusting the model to their particular situation.

There are several issues as regards quantifying the performance of the public sector. Firstly, type of measures needs to be chosen and basic definitions should be agreed. Higher order measures, i.e. measures of efficiency and effectiveness, should rather be preferred over output measures. Concentrating solely on the latter ones will result in merely recording statistics of activity instead of aiming at improving the management and the service of the public sector. Measurement can take place on different levels from micro (individual organisations) to macro-levels (government). If a macro-level measurement is attempted, the method of aggregation from a micro-level may have to be considered and weights should be chosen. If cost-effi-

ciency is compared in the measurement of different public institutions, a comparable cost-accounting system needs to be assured. Furthermore, the results of public sector depend not only on the activity of the sector itself. As various factors play a role, it is proposed to adjust the value of the measures in some cases.

Adequate planning and setting of relevant objectives is a prerequisite to correct performance measurement. On the other hand, an improved setting of objectives is a natural consequence of using the PM in the decision-making process. The objectives set in a specific, measurable, achievable, relevant and timed way will allow for significant measures to be produced. These measures will support managers in monitoring the implementation of the budget as well as its evaluation at the end of the cycle. Performance information can be used in the budgeting process in several different forms (presentational performance budgeting, performance-informed budgeting, direct performance budgeting). In these different concepts the performance information can be linked tightly or loosely to the decision on funding, or used for information only. These concepts have been applied in different public sector fields. The direct performance budgeting is present mostly in education or health sector. It links automatically the performance information to future resources. This particular method is, however, especially vulnerable to gaming. Another purpose of performance measurement is better, more specific planning as well as exact monitoring of results over the period under consideration. Moreover, demand from citizens is increasing for a greater transparency and accountability of the public sector. Internally, through performance measurement all levels of administrative hierarchy are motivated to deliver measurable results.

Not all fields of the public sector can be measured. Additionally, many essential government activities, such as emergency services, must be provided regardless of whether they are used. Moreover, not all inefficient programmes will be eliminated due to their political importance. Also positive outcomes have to be interpreted with caution taking into consideration possible manipulation of data, i.e. gaming. Imprecise objectives as well as unprepared transfer of knowledge for the reform in public sector can have opposite effect to the desirable one. Being aware of the benefits but also possible problems with performance measurement, Poland and other new market economies can profit from this general trend in reforming the public administration funding and assessment.

References

- Ammons D. N. and Rivenbark W. C. (2008), *Factors Influencing the Use of Performance Data to Improve Municipal Services: Evidence from the North Carolina Benchmarking Project*, Public Administration Review, March/April, pp. 304-318.
- Barzelay M. (2001), *The New Public Management: Improving Research and Policy Dialogue*, University of California Press, pp. 156-171.

- Behn R. D. (2003), *Why Measure Performance? Different Purposes Require Different Measures*, Public Administration Review, September/October, vol. 63, no. 5, pp. 586-606.
- Bevan G. and Hood C. (2005), *What is Measured and What Matters: Targets and Gaming in the English Public Health Care System*, The Public Services Programme Discussion Paper Series, no. 0501, December 2005, Economic and Social Research Council.
- Blöndal J. R. (2004), *Issues in Accrual Budgeting*, OECD Journal on Budgeting, vol. 4, no. 1, OECD 2004, pp. 103-119.
- Caiden N. and Caiden G. E. (2006), *Improving Government Performance*, Review of the Book, *High-Performance Government: Structure, Leadership, Incentives*, R. Klitgaard, P.C. Light (eds.), Santa Monica, CA: RAND Corporation, 2005, Public Administration Review, January/February, pp. 139-142.
- Currstine T. (2005), *Performance Information in the Budget Process: Results of the OECD 2005 Questionnaire*, OECD Journal on Budgeting, vol. 5, no. 2, pp. 88-131.
- EC (2000), Commission of the European Communities, *Reforming the Commission*, A White Paper, part I, vol. I, COM(2000) 200 final/2, Brussels, 2000.04.05.
- EC (2004), Commission of the European Communities, *Communication from the Commission. Completing the Reform Mandate: Progress Report and Measures to be Implemented in 2004*, Brussels, .02.10, COM(2004) 93 Final.
- EC (2007b), Commission of the European Communities, *Annual Activity Reports of the European Commission*, http://ec.europa.eu/atwork/synthesis/aar/index_en.htm.
- EC (2007b), European Commission, Directorate-General for Economic and Financial Affairs, site consulted 2008.07.29. http://ec.europa.eu/economy_finance/sg_pact_fiscal_policy/fiscal_policy528_en.htm.
- ECA (2008), European Court of Auditors, *Special Report No 5/2008. The European Union's agencies: Getting results*, adopted by the ECA on 2008.06.05.
- EU (2002), *Council Regulation (EC, Euratom) No 1605/2002 of 25 June 2002 on the Financial Regulation Applicable to the General Budget of the European Communities*, Official Journal of the European Communities, L 248 of 2002.09.16, as last amended by Council Regulation (EC) No 1525/2007 of 17 December 2007 L 343 9 27.12.2007.
- EU (2007), consolidated version of the following documents:
- *Council Regulation (EC, Euratom) No 1605/2002 of 25 June 2002 on the Financial Regulation applicable to the general budget of the European Communities* (Official Journal of the European Communities L 248, 2002.09.16, p. 1),
 - *Council Regulation (EC, Euratom) No 1995/2006 of 13 December 2006*, Official Journal of the European Communities L 390 1 2006.12.30,
 - *Council Regulation (EC) No 1525/2007 of 17 December 2007*, Official Journal of the European Communities L 343 9 27.12.2007.12.27,
 - *Corrigendum*, Official Journal of the European Communities L 25, 2003.01.30, p. 43 (1605/2002),
 - *Corrigendum*, Official Journal of the European Communities L 99, 2007.04.14, p. 18 (1605/2002),
 - *Corrigendum*, Official Journal of the European Communities L 48, 2008.02.22, p. 88 (1995/2006) <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2002R1605:20071227:EN:PDF>.

- Eurostat (2008), Total general government expenditure (% of GDP), overview 1996–2007, (Website consulted 2008.08.05), http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1996,39140985&_dad=portal&_schema=PORTAL&screen=detailref&language=en&product=REF_TB_government_statistics&root=REF_TB_government_statistics/t_gov/t_gov_a/tec00023.
- Hood C. (2007), *The Numbers Game*, Ethos, Issue One, Spring, pp. 32-35.
- IFAC (2008), *International Federation of Accountants*, 2008 IFAC Handbook of International Public Sector Accounting Pronouncements, April, International Federation of Accountants.
- KPRM (2007), Kancelaria Prezesa Rady Ministrów, Departament Budżetu Zadaniowego, *Raport Budżet zadaniowy*, t. I. *Budżet zadaniowy na rok 2008*, Warszawa 2007.
- Lubińska T., Strąk T., Lozano-Platonoff A., Będzieszak M. (2007), *Metodologiczne i wdrożeniowe problemy zastosowania budżetu zadaniowego*, *Ekonomista*, no. 5/2007, pp. 703-716
- Lubińska T. ed. (2007), *Budżet zadaniowy w Polsce. Reorientacja z wydatkowania na zarządzanie pieniędzmi publicznymi*, Warszawa, Difin.
- Lubińska T., Strąk T., Lozano-Platonoff A., Będzieszak M., Godek M. (2008), *Performance Budgeting in Poland*, to be printed in *Revue Française De Finances Publiques*, Paris.
- Melkers J., Willoughby K. (2005), *Models of Performance-Measurement Use in Local Governments: Understanding Budgeting, Communication, and Lasting Effects*, *Public Administration Review*, March/April, vol. 65, no. 2, pp. 180-190.
- Nicholson-Crotty S., Theobald N.A., Nicholson-Crotty J. (2006), *Disparate Measures: Public Managers and Performance-Measurement Strategies Organizational Performance*, *Public Administration Review*, January/February, pp. 101-113.
- OECD (2007a), *Performance Budgeting in OECD Countries*, OECD.
- OECD (2007b), *Towards Better Measurement of Government*, OECD Working Papers on Public Governance 2007/01, OECD Publishing doi: 10.1787/301575636734.
- OECD (2008a), *Economic Glossary*, website consulted 2008.08.05. http://www.oecd.org/glossary/0,3414,en_2649_34109_1968510_1_1_1_1,00.html.
- OECD (2008b), *Glossary of Statistical Terms*, website consulted on 2008.09.04. <http://stats.oecd.org/glossary/detail.asp?ID=1068>.
- OECD (2008c), webpage consulted on 2008.09.04. http://www.oecd.org/countrieslist/0,3351,en_33873108_33844430_1_1_1_1_1,00.html#.
- Pollitt C. (2004), *From There to Here, from Now Till Then: Buying and Borrowing Public Management Reforms*, CLAD Ninth International Conference, Madrid, 3 November, <http://www.clad.org.ve/congreso/pollitt.pdf>.
- Pollitt C., Bouckaert G. (2004), *Public Management Reform. A Comparative Analysis*, Second Edition, Oxford, Oxford University Press.
- Rubenstein R., Schwartz A.E., Stiefel L. (2003), *Better than Raw: A Guide To Measuring Organizational Performance with Adjusted Performance Measures*, *Public Administration Review*, September/October, vol. 63, no. 5, pp. 607-615.
- Van Dooren W., Lonti Z., Sterck M., Bouckaert G. (2008), *Institutional Drivers of Efficiency in the Public Sector*, 36th Session of the Public Governance Committee 16-17 October, paper amended 2008.03.06, Public Governance and Territorial Development Directorate, Public Governance Committee, OECD.

- Weber D. (2007), *Wdrożenie budżetu zadaniowego w Polsce. Ocena postępu, osiągnięć oraz propozycje następnych kroków*, http://www.lubinska.pl/files/materialy/d.webber-opinia_banku_swiatowego.doc.
- World Bank (2008), *Performance-Informed Budgeting in Latin America. Experiences and Opportunities*, Public Sector and Governance Unit, Latin America and Caribbean Region, World Bank, June.
- Yamamoto H. (2003), *New Public Management – Japan’s Practice*, Institute for International Policy Studies (IIPS) Policy Paper 293E, Japan, January.
- Yang K., Holzer M. (2006), *The Performance – Trust Link: Implications for Performance Measurement*, Public Administration Review, January/February, pp. 114-126.
- Yang K., Hsieh J. Y. (2007), *Managerial Effectiveness of Government Performance Measurement: Testing a Middle-Range Model*, Public Administration Review, September/October, pp. 861-879.