COMMENTS ON SESSION 2:
EVALUATING THE EFFICIENCY AND EFFECTS OF PUBLIC SPENDING

Adi Brender[*]

The papers presented in this section cover several aspects of public expenditure (PE) efficiency from different angels. According to the division of labor between the discussants I shall focus on papers 1-3 (Höppner and Kastrop, Herrera and Pang, and Salinas-Jiménez, Pedraja-Chaparro and Smith). However, the third paper discusses the second one thoroughly, a discussion with which I concur, so I chose to focus on the theme emerging from the Höppner and Kastrop paper. Particularly, this paper is favorable to the decision to amend the EU fiscal rules in a way that will accommodate "quality spending". In this discussion I will try to highlight the requirements that the use of the concept of "quality spending" imposes on policy-makers, and argue that we may still be far from being able to use it broadly for policy purposes. Specifically, I will focus on the difficulties in evaluating the efficiency and effects of public spending – a key element in any meaningful definition of "quality spending".

The government's fundamental optimization problem can be presented as:

Max. social welfare (PE) s.t.: cost of revenue collection

where the cost of revenue collection is composed of:

a) forgone private consumption,

b) administration,

c) dead weight loss.

The optimal level of *PE* is achieved when:

$$MU\,(PE) = MC\,(revenue)$$

i.e. when the marginal utility from public expenditure is equal to the marginal cost of increasing revenue collection.

In recent years globalization has been pushing the marginal cost of collecting taxes up, due to intensifying international competition on the location of firms and individuals. Consequently, governments were pressed to adapt by increasing *MU*(*PE*), either by reducing *PE* or by improving its efficiency. Since the competition between governments is not just for low taxation but for the "value for money" that they deliver, the focus is shifting from cutting the overall level of expenditure to increasing its efficiency. Moreover, one should note that while increasing efficiency improves welfare it does not necessarily reduce the level of public expenditure. Such a reduction is likely to take place if the efficiency gains occur in activities that are far from the margin (necessities), because in these cases

[*] Research Department, Bank of Israel. E-mail: adib@boi.gov.il

only the income effect applies.[1] However, if the efficiency gains affect marginal activities (those that are less critical), both the substitution and income effects apply and the result may be an increase in public expenditure.

A decision to adapt the fiscal criteria to allow more room for "quality expenditure" requires that there would be a clear definition of what "quality expenditure" is and a system for the evaluation of various expenditures according to their quality. In the remainder of my comments I will depict what may be required for an evaluation of expenditures according to their quality, and what are the merits and risks of such an approach.

A key first step in a meaningful evaluation of *PE* is defining government objectives. These objectives should reflect the needs of the individual country and the tastes of its population. Clearly, a government should know what it wants to achieve to be able to evaluate whether a certain expenditure is useful or not. Based on these objectives, policy-makers should examine their programs and assign to them properly set (social) values**.**

If the government is able to set appropriate values for various programs and areas of activity, the next step is to adopt reliable methods for evaluating performance and measuring output. Such methods can contribute substantially to improving the quality of *PE* by identifying weaknesses and by facilitating the adoption of performance-based remuneration and/or budget allocations through the identification of success in reaching pre-specified targets (preferably as part of a long-term plan). More specifically, if the evaluation system properly measures the desired outputs, identifies their respective cost, and the outputs are properly related to outcomes, then the system indicates the "value for money" generated by *PE*, and may be useful in pointing to potential welfare-enhancing reallocations.

Unfortunately, the conditions mentioned above are quite demanding and require significant investment in the design, collection and processing of information. This type of evaluations is particularly useful where outputs – and outcomes – are easy to measure. However, this is the exception in the public sector; in most of its activities outputs are not easily identifiable and even more rarely are they measurable in a meaningful way.

Performance can be evaluated according to different measures, and one should clarify – in advance – which measure is to be used when defining the quality of expenditure:

- technical inefficiency: output is lower than the possible maximum – given the existing inputs (x-inefficiency);
- economic inefficiency: the input composition could be changed to produce more output at the current cost (e.g., by shifting to highly paid high quality staff);

---

[1] The income effect works to reduce public expenditure because the savings from the efficiency gains (the lower cost of producing the original level of services) will be reallocated between public and private consumption according to the population's tastes.

- technical ineffectiveness: the output is not useful in generating the desired outcome (e.g., hospitals may be much less effective than immunization in containing certain diseases);
- economic ineffectiveness: reallocating resources to activities that produce other desired outcomes may increase overall welfare at the same cost (e.g., building safer roads may be more effective than improving the health system in reducing mortality).

Beyond the question of which definition of efficiency or effectiveness underlies the definition of "quality spending" one should also decide which performance criteria are going to be set and on what targets will they be based, as these may vary considerably:

- outputs or outcomes. In most areas only outputs can be measured. However, outputs may not always provide sufficient information to evaluate effectiveness. For example, test results may not be a good enough indicator for the quality of education achieved by the school system;
- minimizing the cost of specific inputs by constraining their use or focusing on minimizing overall costs while allowing flexibility in the choice of inputs. The choice between these two options would depend on whether policy-makers are focusing on predictability and discipline, or on economic efficiency;
- monitoring outputs or inputs: Should entities' evaluations be based on their ability to produce a given range of outputs, using whatever inputs they wish (minimizing cost), or on their ability to produce outputs given a fixed set of inputs (maximizing output).

The choice between these criteria may be a difficult task even in a private firm. Clearly, identifying a meaningful criterion to be used for a cross-country pact, and being able to measure it, is a very complicated process.

Once the criterion/criteria for evaluating performance had been set, one needs to decide according to which dimensions will performance be evaluated. There are various dimensions that could be used such as past performance, plan, a predetermined standard or peer comparison. Policy-makers should clarify which concept, criterion and dimension are more relevant to each unit and each type of expenditure, otherwise evaluations may turn into a source of confusion. According to the concept used in the UK targets should be set in a SMART way (that is: Specific, Measurable, Achievable, Relevant and Timed) to be effective. These criteria may be a good guideline for evaluating the programs that qualify as "quality expenditure" once the other conditions mentioned above are also met.

Despite its substantial merits, the adoption of evaluation-based management and remuneration system may create some risks as well. First, units may focus on achieving the pre-set outputs while ignoring other, which are as important for achieving the desired outcome. Second, the government may find itself focusing on measurable, rather than important, criteria. Finally, lifting budgetary controls, while relying on evaluations, may hurt efficiency and budgetary discipline if the targets

are not properly set, and the evaluation/quality of information is not carefully monitored and verified.

Another risk to the efficiency and effectiveness of public expenditure, due to the focus on measurable quantitative outputs, is related to the quality of service and output. Since quality in public services is hard to measure there may be a tendency of units to curb quality in order to perform better on the quantitative scales. Theoretically, one could use in the evaluation process quality measures as well but, in practice, both internal and client quality valuations have substantial deficiencies.[2]

The usefulness of performance evaluations is also hampered by data and technical difficulties. To be able to improve the reallocation of public resources between activities the evaluation should provide information on the marginal output and efficiency of the various activities. However, in practice, only data on average efficiency are available, and these may not be indicative for efficient decisions. Additionally, a significant part of the public sector deals with prevention of undesired outcomes (defense, fire protection) and the measurement of efficiency in these areas is particularly difficult.

The Data Envelopment Analysis (DEA) discussed in three of the papers in this session may be a useful, though limited, technique for evaluating the efficiency of public expenditure. Notwithstanding the methodological issues raised by Salinas-Jiménez, Pedraja-Chaparro and Smith, this method deals with technical and economic efficiency and can be used for performance benchmarks in parallel units where output is measurable (e.g., local authorities, international). However, the use of this method suffers from the problem of output diversion which may be quite substantial.[3] Additionally, one should be careful in interpreting the results of DEA. For example, the findings by Herrera and Pang that the efficiency of *PE* is decreasing when the share of *PE* in GDP is rising do not necessarily imply that *PE* should be reduced; if social preferences are for more *PE*, then the increasing marginal cost is a price that society is willing to pay for public goods. Only if one can show that the marginal cost of *PE* exceeds the marginal value of that expenditure can these results be used as an indication for the need to reduce *PE*. Unfortunately, our ability to make such judgments is still limited.

---

[2] Although using client response may be more useful for generating decision-makers' (politicians') interest in quality performance.

[3] For example, teaching students for international tests rather than improving education systems.