Output, Input and Incentives Measurement for Non-market Activity
The case of Italian Revenue Agency

Marco CUTAIA(*) and Stefano PISANI(**)

(*) University of Milan - m.cutaia@libero.it
(**) Italian Revenue Agency - stefano.pisani@agenziaentrate.it

1. Introduction

The *Italian Revenue Agency* (hereinafter Agency) is a new Government body, founded in 2001, and it has the strategic objective of maintaining and encouraging taxpayers in their willingness to comply with their tax duties as laid down by law. It works in a self-government framework, regulated by an annual agreement with the Ministry of Economy, who sets the aims to be reached and the related budget. The Agency has two main goals: improving the quality of provided services and the efficiency in its production process. In order to achieve these two goals, analytic tools are now needed to monitor the input, output and productivity over time.

One possible tool to improve efficiency is an incentive scheme based on performance and effort. In the Italian case we have a multi-tier relation: at upper level the Ministry of Economy bargains with the Agency on the remuneration to be paid out and the objectives to be reached; at lower level the Agency negotiates with its offices over the distribution of reward.

A principal-agent analysis applied at the relationship between the Ministry and the Agency allows us to obtain an efficient solution and the consistency between the two steps of the bargaining. The principal (the Ministry at upper lever, the Agency at lower level) tries to maximize his expected utility offering a compensation scheme that serves a dual function: allocating risks and rewarding productive work. The agent (the Agency and its offices respectively) maximizes his expected utility as well, that includes some personal concerns for output (such as public service motivation or career concerns). In this paper we will focus only on the first tier of the model, where the Ministry and the Agency as a whole are parts of the principal-agent relationship.

The theory of incentives generally assumes that outcome of a productive unit are verifiable but its actions are not. This kind of productive units are defined as craft organizations (Wilson's, 1989). In the case of an organization that provides non-market services (like a government Agency) the problem is more complex, because, neither outcomes nor actions are easily observable. For this reason, in such Agencies, explicit incentives are very weak, and there will be high conflict between management and lower-tier operators. As a consequence of this measurement problems, the management concentrates his attention on the more easily observable dimensions and denies the operators much freedom of action; the operators will engage in the immediate tasks they regard as essential while keeping the management satisfied about his focus (Dixit 2002).

Considering these measurement problems, it is extremely difficult to use an incentive scheme to improve both the quality of General Government outcome and the efficiency in productivity process of non-market activities. These difficulties in evaluation arise not only from the peculiar characteristic, typical of all services (e.g. intangibility), but also because in the case of General Government it is impossible to determine the output value simply multiplying quantity by price, because non-market services do not have a price or it is not economically significant (Griliches 1992, Jorgenson and Fraumeni 1992, Murray 1992).

This paper suggests the application at the micro level (the Agency) of some tools derived from the national account system (SNA 1993, Eurostat 2001). In particular the best practices to evaluate the output at constant prices have been applied. In the mean time, matters related to the quality aspects of services have been analysed and discussed. According to the National Accounts definition, the proposed indicator measure only the efficiency gains (or loses) achieved by the Agency, excluding any implication on the effectiveness (outcome) on the provided services.

The theoretical approach to measure the efficiency improvement is based on the monitoring of the dynamic of the apparent labour productivity. It is well known that this measure is obtained by comparing the output at constant prices with the input of labour. In this paper the input of labour has been measured as labour input at constant compensation, in terms of hours actually worked (see OECD, 2002).

1

In the second paragraph we analyse the incentive scheme actually used in the public sector, then we introduce a simplified principal-agent model. The fourth paragraph considers the extension of the principal-agent model to the multi-task case. In the following we illustrate the tools to use in order to obtain reliable measures of output, input and productivity. Finally some conclusions are drawn.

2. Incentives schemes in the public sector

Many government agencies use a budget based reward system, where the reward (annual bonus) is calculated on the basis of the comparison of objectives (performance standards) and results (performance measures). As shown in Figure 1, the pay-performance relation is the function that links results and payments, and it could vary in the incentive zone from a minimum, at the performance threshold, to a maximum, at the bonus cap level, beyond that any other production improvement doesn't cause a reward raising.

The Italian Agency uses a reward scheme as that depicted in figure 1, limited to few easily measurable production activities. The budget based incentive system is ineffective to promote productive work, because:
- It is referred to few output dimensions and determines an effort reallocation from the harder to measure activities to those easily measurable, for substitute actions (Holmstrom and Milgrom, 1991);
- It causes inefficient agent's behaviours to influence the budget process.

The aim of this work is to propose a new performance measurement system based on the overall evaluation of the Agency's output.
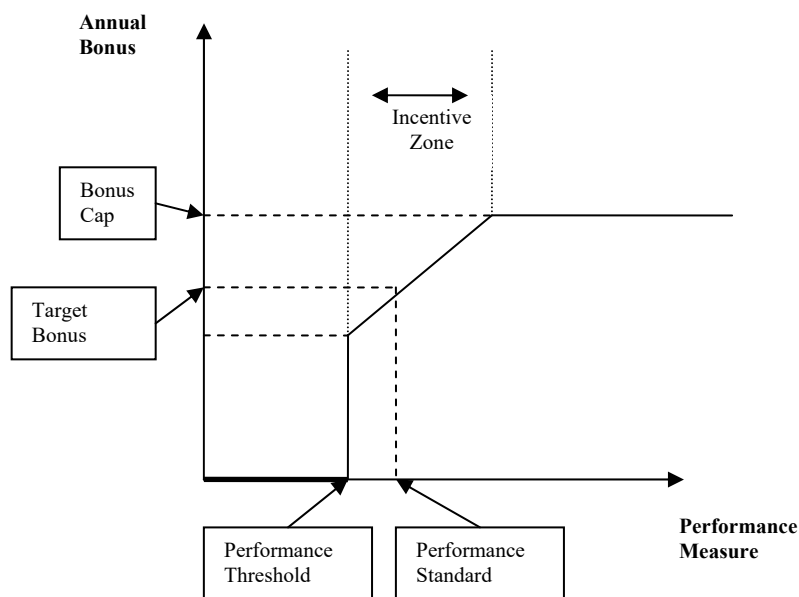

Figure 1 – Typical incentive plan components (Murphy, 2000)

2.1 Performance standard

The main economic problem to solve in the determination of an incentive scheme is related to the construction process of performance standards and reward (Murphy, 2000). Such process could reflect mainly the agent's action or could be conditioned by external constraints. Murphy found that the budget setting up processes of the big enterprises is internally determined (89% of the analysed sample).

Internally determined standards could be set on peer-groups performance, on past performance, or on a *a priori* complexity evaluation of the activities. Each methodology is characterised by weaknesses that are summarised below:

2

Peer-group performance standards cause free-riding into the organization: agent has the incentive to sabotage co-workers or collusively shirking. The free riding could be faced by relaxing the budged-balancing constraint saving the surplus that exceed the achievement of the targets (Holmstrom, 1982).

If the prior-year budget and performance standard (such as the ratcheted target) is used, agent avoids actions that could raise the target in future years[1] and it causes an income smoothing effect. Once the agent reaches the performance standard, he stops any further effort, in spite he could produce more. The budget based on prior-year performance could not reflect the firm's real capabilities and the market opportunities.

The *a priori* complexity evaluation method is characterized by an high degree of subjectivity and could be affected by the agents gaming behaviour.

At the present, the method followed by Italian Agency is a mixture of peer group performance and a priori complexity evaluation. In this paper we suggest to enlarge such criteria in order to stimulate the rate of growth of output and to reduce the uses of human resources.

2.2 Performance measures

In many cases Government Agencies fail in the definition of means necessary to stipulate an incentive contract (Baker, 1992). In this case it raises the problem of determining performance measures consistent with the "real" target function of the principal.

The strategic behaviour of the agent could reduce the power of incentives, that is, agents could influence the performance measure without affecting their real output. In order to avoid this dysfunctional situation high correlation between marginal product on performance measure and marginal product on principal's objective function is needed. If this correlation exists a piece-rate contract is efficient.

Ittner and Larker (2002) affirm that the operational measures (i.e. productivity), as well as behavioural ones, are more effective than accounting measures. The method proposed in the paper, based on the production at constant price calculation, lies in these suggested categories.

In order to calculate the production at constant price, in a multitask environment, is crucial to determinate in the right way the weights used to aggregate different performance indicators. The use of subjective weights permits the principal to face the agent's gaming on the compensation system (Baker, Gibbons e Murphy, 1994). As we will show in the next paragraphs, the use of weights derived from the costs structure in the base year (or the strategic structure) is one of the possible solutions to share the effort allocation between principal and agent. In this context, a fixed proportion among the tasks of the agent is imposed. This proportion is equal to the relative impact of tasks on performance measure and to the planned strategic allocation of the organization. The technology used to realize the performance measure allows us to reach a set of feasible actions that minimize the agent's personal cost at any level of performance[2].

The agent's optimal choices between tasks T_1 and T_2 are shown in Figure 2, given a performance measurement system linked with an incentive scheme. Following this model the marginal rate of substitution between the tasks is constant at any level of production, as well as the implicit proportion on incentives. Given a reward level the agent optimally chooses the couple of actions that minimize her private cost to exert effort, lying on the line A. The isocost curves reflect the increasing marginal cost hypothesis. The model suggests, on the one hand, that point **a^** is the sub-optimal solution (at a lower production level) affected by noise performance measure, under the hypothesis that the agent is risk averse. On the other hand, point **a^^** is the second best optimal solution with precise performance measure but characterised by incongruity between the performance measure and the principal's objective function.

In this paper, a single performance indicator has been calculated for each task (i. e. T_1 and T_2). Such indicators have been summarised using a weighting system that could reflects the trade off between the tasks. According to figure 2, the obtained result lies on line A. The degree of optimality of the achieved solution depends on the congruity and on the noise that could affect the performance measures.

---

[1] Murphy Kevin J. (2000).
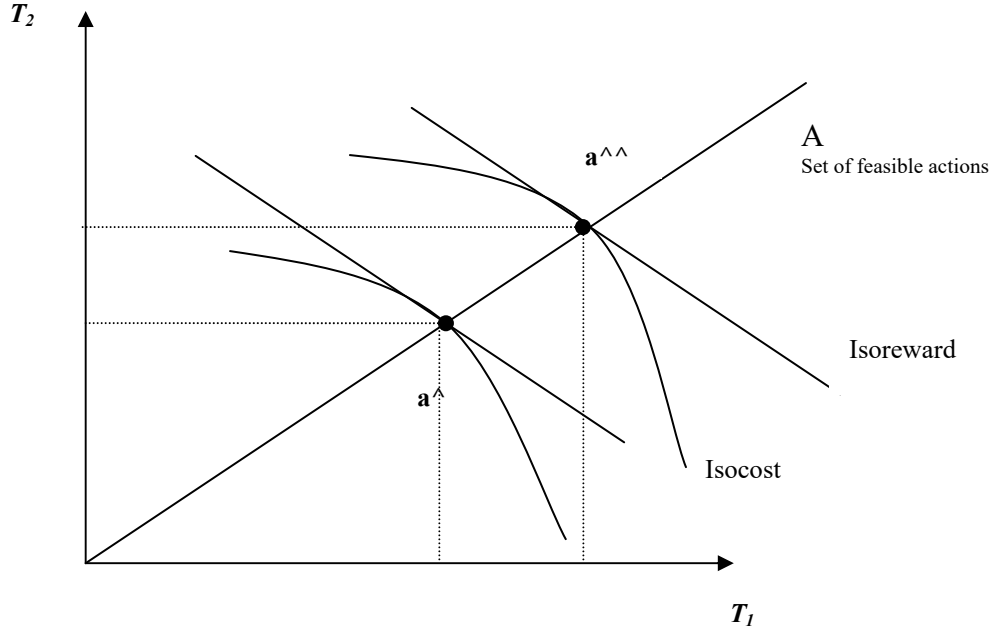[2] Feltham, Xie (1994).

Figure 2 – Set of feasible actions, given a performance measurement system.

3. A simplified principal-agent model

The principal-agent relationship occurs between two subjects: the principal that intends to carry out a plan and to do this hires a manager (the agent) that provides for an effort to achieve the targets scheduled in the plan. In this paper we use a simplified version of the model proposed by Holmstrom and Milgrom (1991). In particular we reduce the multi-task model to a single task case, by using a single performance measure.

The principal maximizes his utility ensuring the agent at least the same utility level he could obtain out of the relation, and he should implement exactly the level of effort chosen by the agent.

The principal is an expected utility maximizer, and the utility function (U_p) is assumed to take the form

$$U_p(CE_p) = E\left\{-\exp\left[-R_p\left[B_p(x) - w(x)\right]\right]\right\}$$

where: (CE_p) is the principal's certainty equivalent to the uncertain situation (lottery) consisting of the expected benefit (B_p) minus the wage (or payment scheme, w) multiplied by a risk premium (R_p), under the hypothesis of constant absolute risk aversion; (w) is a function of the output that could be produced by the agent *(x)*.
In the main time, the agent maximizes the certain equivalent value (CE_a) derived through his utility function (U_a)

$$U_a(CE_a) = E\left\{-\exp\left[-r_a\left[B_a(x) + w(x) - C(t)\right]\right]\right\}$$

where $B_a(x)$ is the agent's private valuation of output and $C(t)$ is the cost of effort. The certainty equivalent income then consists of private benefit plus the wage minus the cost of effort and a risk premium. The effort generates an information signal, which could be thought as a production function

$$x = \mu(t) + \varepsilon,$$

where ε is stochastic term normally distributed with mean zero and variance Ω.

4

As Holmstrom and Milgrom (1987) demonstrate, we could restrict our attention to linear payment schemes; as such formulation is actually used in many Government Agencies[3],

$$w(x) = \alpha x + \beta$$

where α is the incentive component and β serves only to allocate the total certainty equivalent income between the two parties, and to meet the agent's participation constraint.

The first best solution with observable effort, obtained maximizing the total surplus, is

$$\left\lfloor B_p{}'(\mu(t)) + B_a{}'(\mu(t)) \right\rfloor \mu'(t) = C'(t)$$

where ' indicates the partial derivative by the effort (*t*).

When $B_p{}'(\mu(t)) = \alpha$, the 100% powered incentives are achieved. In this case the verifiable effort allows the principal to implement the desired productive solution up to the principal's marginal cost equals his marginal benefit.

Figure 3 shows the equilibrium between principal and agent in terms of payment scheme (*w*) and output (*x*). According to Holmstrom , Milgrom (1991) and Dixit (1997), we assume a quadratic form for the cost of effort, that explains the concavity of the agent's indifference curve. Point A represents a first best solution because it lies on the highest reachable principal (U_p) and agent's (U_a) indifference curves, given the compensation scheme (*W*). Any other point of figure 3 represent a sub-optimal solution; for example, if the principal would implement a larger level of output (point B), he would cause a loss of welfare and he reaches a lower level of utility. The fact is that, in order to reach the point B the principal has to pay to the agent a compensation that is greater than his marginal benefit. In the main time, by virtue of the participation constraint, in B the agent benefits from the same utility level of A.
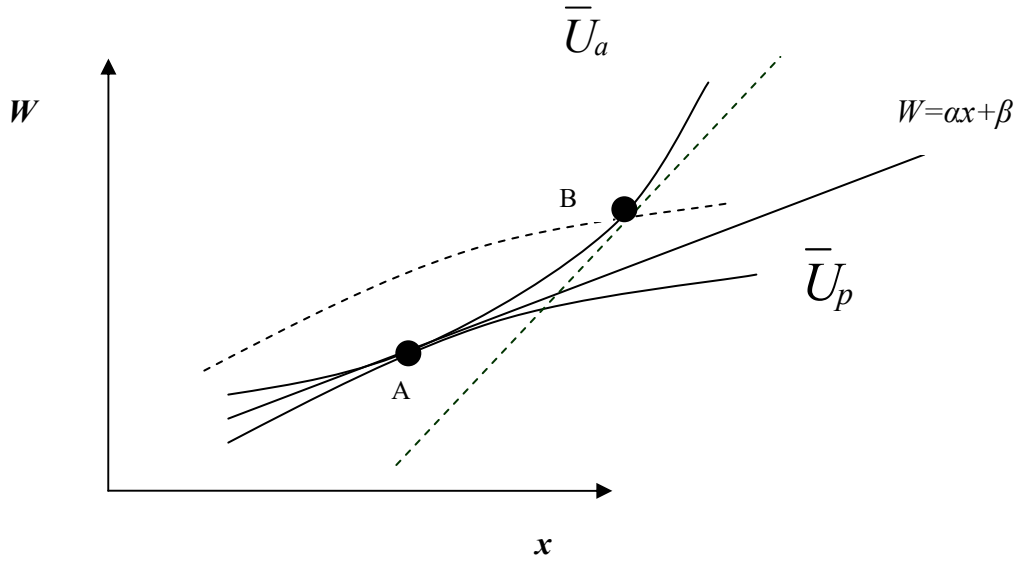
Figure 3 – First best equilibrium (point A) and sub-optimal equilibrium (point B).

Output and input measurability problems, even with agent's risk aversion, induce a sub-optimal solution. In order to avoid this eventuality, in this paper we suggest to use techniques deriving through the system of national accounts. It is necessary to stress that the necessary conditions of a first best solution, in a multitask environment, are the agent's risk neutrality and the measurement precision. Furthermore, if the principal's objective function is not contractible, as underlined by Baker (1992), a perfect congruity between the marginal product on the performance measure and on the principal's actual objective function is necessary for first best.

As shown in figure 2, the error in the measurement system determines a sub-optimal effort intensity and an equilibrium at lower levels of reward, production and general welfare.

---

[3] Restriction to linear payment schemes can be justified if the model is the reduced form of a dynamic one where the agent controls the drift of a Brownian motion.

5

We must remark that the crowding out of non-measurable activities (i.e. quality) by the measurable ones, occurs when tasks are substitute in the agent's cost function and the incentives are given only on the better-measurable tasks (Holmstrom and Milgrom, 1991). Crowding out doesn't occurs if the tasks are complementary in the agent's cost function.

4. Multitask considerations

We must remark that the Agency, as well as other government agencies, performs several tasks, and the interactions among these dimensions, whether they are substitute or complement, are not clarified in advance. This multidimensionality of goals goes hand in hand with two difficulties: Several dimensions of performance are hard to be measured; and it is very difficult putting weights on each dimensions for specify a government welfare function (Tirole 1994). Furthermore, the government welfare function is often not contractible.

Holmstrom and Milgrom (1991) obtain these remarkable results: (i) The desirableness of providing incentives for one activity decreases with the difficulty of measuring performance in the other activities that make competing demands on the agent's time and attention, so an optimal incentive contract could consist in paying a fixed wage independent of measured performance; (ii) separating tasks according to their measurability characteristics allows the principal to give strong incentives for tasks that are easy to measure without fearing that the agent will substitute effort away from other harder-to-measure tasks. When possible, Agencies should be organized to group together complementary activities, and grouping of substitute activities should be avoided (Dixit 2002).
The job-design considerations by Holmstrom and Milgrom (1991) cannot always be applied. Synergies between activities do not allow the specification of several different contracts and job groups.

We'll show in the next paragraph how to obtain a single output indicator. This is a way to put weights on any one activity the Agency provides, and we restrict the agent and the principal to share this vector of weights, representing the strategically optimal allocation of effort. Our results show that we obtain the same output and we have no efficiency losses; we reduce the variance of the error term[4], so the risk adverse agent strictly prefers the new equilibrium at a higher utility level.

In the Figure 4 it is shown that a compensation scheme based on the synthetic output indicator determines a new equilibrium with a higher incentive component and higher utility by the principal. The principal can guarantee the reservation utility level to the agent, using new resources to increase the production level and the corresponding reward at a higher principal's utility level (from point A to point B in figure 4).
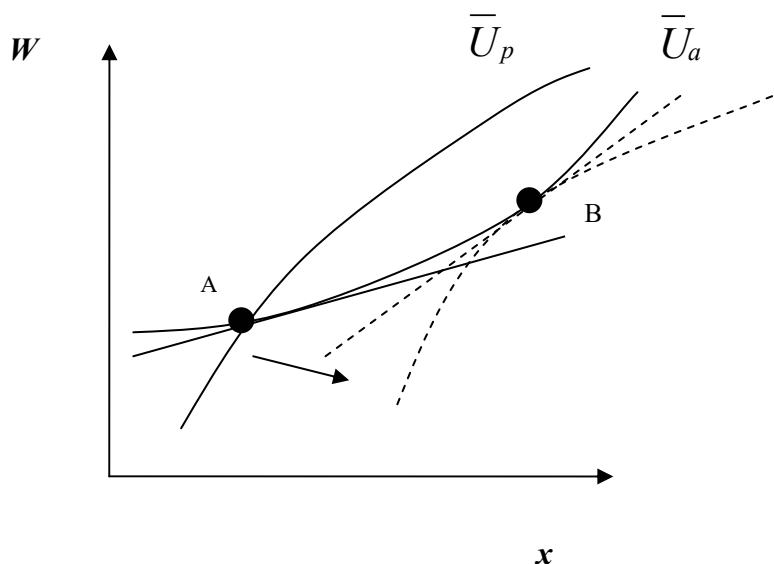

Figure 4 – Aggregate reward and welfare improving.

---

[4] The reproductive property of independent normal variables applies.

A reasonable way to choose the weights to be used is demanding it to the principal, especially to promote organizational changes. As future development of the research, the weighting scheme based on the budget of the Agency (table 1), could be replaced by an "optimal" weighting scheme. Such scheme could be based on the opinion expressed by a group of "decision makers" (e.g. the political authorities at the upper level of the bargaining and the top management of the Agency at the lower level). Within the decision group, the opinions of members could be elicited by means of a careful technique that harmonize their different point of view (e.g. the Delphi method).

The use of this system of performance measures and incentives could cause the first best optimum only incidentally, if the performance measurement system reflects the relative effort provided to achieve each task, and the relative effort is equal to the task's relative contribution on the expected payoff of the principal. Figure 5 is a modified version of figure 2, including the first best solution. A is the set of agent's optimal response, given the weighting system, a^* is the first best optimal allocation and production, $a^{\wedge\wedge}$ is the second best optimal action with performance measure precision and non congruity and $a^{\wedge}$ when this measure is noise (that is it has a high variance) or when the agent is risk averse (Feltham, Xie, 1994). Point a^* could be reached only if there is a perfect congruity between the performance measure and the principal's expected payoff. If it is not, there will be a dysfunctional incentive system that causes a inefficient effort allocation ($a^{\wedge\wedge}$).
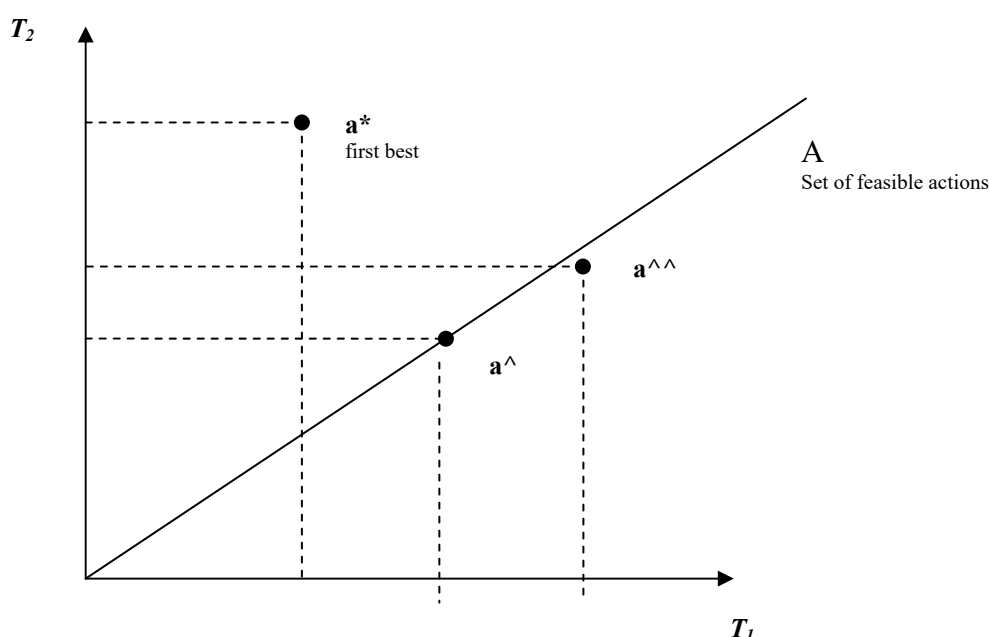
Figure 5 – Impact of noncongruity and imprecision of a single performance measure on the optimal effort allocation among tasks (Feltham, Xie, 1994).

The Agency bargains with the Ministry of Economy on few activities among several services provided. This fact, even with the agent's risk aversion and substitutability of tasks in the cost function, determines optimal weak incentives and a second best equilibrium. In this context public service motivation and personal concerns play an important role. But if some activity is not measurable at all, a generalized incentive fall occurs, because of the agent's optimizing behaviour (Dixit, 1997). Furthermore, the Agency can not apply the job design rules (Holmstrom and Milgrom, 1991), because many activities generate positive externalities due to contemporaneous provision.

5. Output and input measurement[5]

Faced with growing budget deficit and increasing demand on fiscal services, Ministry of Finance drafts an agreement with the Agency, in order to fix a compensation to remunerate a basket of services that must be offered by the Agency. The compensation is split in two components: the first one is devoted to the remuneration of the ordinary functioning and the second is an incentive scheme bound by the achievement of specific targets. The targets are expressed in terms of volume of output of the four "macro-activities" of the Agency: 1) Tax System Management; 2) Legal Advice, 3) Tax Assessment; 4) Customer Services.

---

[5] The terminology used in this paragraph conforms to the international standard used for performance measurement. See OECD (2000, 2001), Eurostat (2001), UN et al. (1993).

5.1 Output measurement

In order to have a correct measurement of Agency output, it is necessary to define carefully the object of measurement, i.e. the produced services. To this end it is useful to make reference to T. P. Hill's definition, by which services are outputs produced to order and typically consist of changes in the conditions of the consuming unit realized by the activity of producers at the demand of the consumers. The definition underlines the central role played by the consumer unit in considering the services output; for this reason the output of the Agency covers only the services provided to external users (consumers), and only those, and excludes activities that are in fact ancillary[6] to the former. Such external services are included in the "macro-activities" that represent the target of the Agency.

Each "macro-activity" includes a basket of different services (named "products") extremely heterogeneous among themselves. The problems of measuring price and volumes for these products arise from the fact that, by definition, no market price exists. In order to solve this problem, one will have to rely on volume indicator method; in particular, we use indicators related to the unit of output, defined as an actual delivery that takes place from the producer (the Agency) and the consumers (the taxpayers). As reported earlier, the proposed method is consentient with SNA (1993) that recommends that the most appropriate way to measure government output is by using direct volumes measures[7]. In particular, we adopt the guidelines suggested by Eurostat (2001, par. 3.1.2.1). Such guidelines state that, in order to define as appropriate an output indicator, they should:
1. covers all services produced by the producer that are provided to external users, and only those; activity that are in fact ancillary to the main output should not be counted;
2. be weighted by the cost of each type of output in the base year;
3. be defined as detailed as possible;
4. be quality adjusted.

As mention before, considering only the 4 macro-activities, which represent the target of the Agency, satisfies condition 1. Furthermore, in order to meet condition 3, the adopted procedure assigns a single volume indicator (i.e. the fiscal audit for the small enterprises) to each "product" at a very detailed level of activity (137 items, see annex 1). The indicators used in the calculation of the output are the same that are included in the annual agreement between the Ministry of Economy and the Agency. Such indicators are related to the efficiency of the production process and they contribute to determine the incentive scheme of the Agency[8].

In order to summing up a volume indicator by "macro-activity" we have to face the problems related both to the different unit of measurement of each volume indicator and to the different complexity (quality) that characterize each "product" (point 4 of Eurostat recommendation). A weighting system based on the "unit average time of work" (UTW) allows us to distinguish between the basic and the complex activities and to transform all the indicators in a homogeneous unit of measurement, expressed in terms of time equivalent of work, (TEW), that represent volume indicators for the "macro-activities". The UTW is not calculated on the basis of the actual time spent to accomplish a task, but is referred to an "ideal" efficient production process. Using an ideal UTW we adjust for the quality each indicator (point 4 of Eurostat), which means the greater the UTW is the higher the quality of the provided service is.

Given the shortness of time series, the UTW is fixed in the base year (in our case the 2001) and then applied to all the examined period (2000-2002). In the future, when we will conduct an analysis on a longer period, we will change the basis every year and we will use a chained index. In this case, in order to stimulate the increase in efficiency, for the same activity, the UTW of a following basis must less or at least equal to the previous.

In order to calculate the total output according to point 2 of Eurostat recommendation, the volume indicator of each "macro-activity" has been summed up using as weighting scheme the cost share derived from the budget of the Agency

---

[6] An ancillary activity is a supporting activity undertaken within an enterprise in order to create the conditions within which the principal or secondary activities can be carried out; ancillary activities generally produce services that are commonly found as inputs into almost any kind of productive activity and tha value of an individual ancillary activity's output is likely to be small compared with the other activities of the enterprise (SNA 93 par. 5.9 and 5.10).
[7] The Office for National Statistics of United Kingdom has conducted a survey on 43 countries asking about progress made in using direct volume measures to obtain government output. The results show that the most exstensive use is UK and New Zealand (60%-70%), followed by the Netherlands, Italy and Australia (20%-50%). Jenkinson (2003).
[8] The overall incentive scheme includes also an indicator related to the outcome of the Agency (namely the sum collected through the tax audit). At this stage such indicator is ignored.

8

in the base year (see the column a of table 1)[9]. As well as in case of UTW, when the analysis will be applied on a longer period we will leave the fixed base to adopt a chained index system.

Table 1. Main economic indicators of the Agency – growth rates, years 2000-2002

Macro-Activity	Budget weighting scheme 2001	2000-2001			2001-2002		
		Volumes Indicators and production at constant price	Input of labour	Apparent labour productivity	Volumes Indicators and production at constant price	Input of labour	Apparent labour productivity
1) Tax System management	0.48	11.51%	1.51%	9.85%	-2.63%	-6.91%	4.60%
2) Legal Advice	0.02	111.13%	107.45%	1.77%	-12.90%	-14.07%	1.36%
3) Tax Assessment;	0.31	2.47%	-3.46%	6.15%	6.54%	7.23%	-0.64%
4) Customer Services	0.19	-5.53%	-10.88%	6.00%	-5.22%	-6.11%	0.95%
Total	*1.00*	*5.65%*	*-3.26%*	*9.21%*	*-0.41%*	*-0.82%*	*0.42%*

The weighting scheme adopted is equivalent to calculate the overall output of the Agency through the aggregation of the volumes indicators using the Laspeyres formula (last rows of columns b and e of table 1). In this way a measure comparable with the national accounts output at constant price is obtained. Furthermore, diving the value index, expressed through the rate of change of the budget of Agency, by the rate of change of output at constant price we can calculate a price indicator expressed in terms of implicit deflator (Paasche index).

5.2 Input measurement

The main input of the Agency is Labour, and the total hours actually worked (THAW) is the preferred measure of labour inputs. THAW is a physical measure of labour input. It is based on the assumption that the mix of different kinds of jobs is much the same at different time examined. Labour input reflects the time, effort and skills of the work force. While data on hours worked captures the time dimension, it does not reflect the skill dimension. Hours of highly skilled persons and hours worked by unskilled persons cannot simply be added to obtain an aggregate measure of labour input; they have to be weighted by their respective relative productivity to account for differences in skills. A possible approach to capture the differences in the productivity is given by the indicator of the total hours actually worked at constant compensation (THAWCP), that is what the employee's current annual compensation would be if it were calculated using the level of compensation ruling during a selected base period (2001)[10]. In particular, nine different professional skills have been considered to calculate THAWCP, which are representative of the different "quality" of input of labour involved in the Agency productive process.

An example of the THAWCP calculation procedure is shown in table 2, where the employee's compensation[11] reflects differences in marginal products among workers. In other words, the quality of labour input is defined as the ratio of labours input (THAWCP) to hours worked (THAW). A measure of total input of labour (quality adjusted) can be derived by adding THAWCP across all 8 categories of labour input of the Agency. Changes in the quality of hours worked represent the differences between changes in the weighting quantity index (-3.26% in table 2) and changes in un- weighting index of hours worked (-2.47%).

The weighting scheme adopted in this paper is a simplified method to capture the quality of input of labour. It could be applied in the short-term analysis. In the log period, if the weights are held constant, annual changes in marginal products are not reflected accurately. The resulting estimates of year-to-year productivity change are biased. In this

---

[9] The same methodology should be applied also to obtain the volume indicators of the macro activity. At the present it is not possible to follow such procedure because a system of book-keeping so detailed is not available.

[10] See UN et al. (1995, par. 17.11-17.32). This methodology is founded on the theory of the firms that, under certain conditions (the forms is a price-taker on labour markets and aims at minimising its total costs), labour of a certain type will be hired up to the point where the cost of an additional hour of labour is just equal to the additional revenue that this labour generates.

[11] The cost of labour input from the point of view of producer also includes supplements; differences in wage and salary income incorporate also employers' contribution to Social Security and other supplements to wages and salaries. Jorgenson (1990, pp. 35-36).

9

case, more general model should be applied that express labour input as a translog function of its individual components (Jorgenson 1990).

Table 2. Hours actually worked and hours actually worked at constant compensation (Years 2000 and 2001)

Professional skills	Hours actually worked THAW (x 1,000)		Employee's compensation (per hour) It liras	Hours actually worked at constant compensation THAWCP		Rate of change of THAW and THAWCP (2001-2000)	
	2000	2001	2001	2000	2001	THAW	THAWCP
	(1)	(2)	(3)	(4)=(1)x(3)	(5)=(2)x**(6)**	(7)	
A1	2,510	2,564	11,744	29,477	30,112	2.15%	2.15%
B1	2,238	2,267	12,245	27,404	27,759	1.30%	1.30%
B2	20,697	19,120	12,918	267,364	246,992	-7.62%	-7.62%
B3	5,683	6,587	14,199	80,693	93,529	15.91%	15.91%
C1	14,269	13,194	15,343	218,929	202,436	-7.53%	-7.53%
C2	2,105	2,659	16,629	35,004	44,217	26.32%	26.32%
C3	4,141	4,147	18,203	75,379	75,488	0.14%	0.14%
Managers	1,669	1,459	65,802	109,824	96,005	-12.58%	-12.58%
Total	53,312	51,997		844,074	816,537	-2.47%	-3.26%

The THAWCP has fallen both in 2001 and in 2002 (table 1), this is mainly due to the reduction of number of employees subsequent to the re-organization of the fiscal administration (realized in 2000) and the creation of a new organizational model (the Agency) characterized by a more stringent accountability mechanisms. The reduction has been realized in a selective way according to the kind of activity performed by the employees. In fact, the internal activity has been reduced more than the external (respectively –7,16% and –2,46%). This is coherent with the modernising Agency programme that has trying to set out a new vision for public service that meets the needs of citizen, not the convenience of services providers.

5.3 Performance and effort in a principal-agent model

Finally, the production at constant price and THAWCP represent two valid indicators to approximate, respectively, the output and the effort provided by the agent in the principal-agent model.
So it is possible to use the production at constant prices as a performance measure in a reward system, with the caveat that performance measures frequently are incomplete or imperfect representation of the economic consequences of the agent's actions (Feltham e Xie, 1994).

Each product measure is naturally characterised by a proper variance. The aggregation process, under the not stringent hypothesis of normal distribution of probability, reduce the overall error variance, because weights are each less than one and the reproductive property of the normal variable applies. It is possible to ordinal estimate the variance of the error terms and to use this estimation in the reward system building.

Given that the measure of output at constant price is independent by the measure of input of labour, we can easily calculate a raw measure of the apparent labour productivity[12], as a ratio between output and input. By construction the productivity showed in table 1 measures the improvement in efficiency in the production of output and doesn't consider the progress in the effectiveness (outcome).

This measure is affected by another drawback, due the fact that gross output represents the value of production without correcting for purchases of intermediate inputs. In order to consider this phenomenon it is necessary moves from the evaluation of gross output to net output (or value added, see OECD, 2000, par 3.1.3). At the present it is possible to perform such evaluation only for the total of Agency but it is not computable for each macro-activity. In order to ensure coherence between the figures of each macro-activity and the overall Agency the gross output has been chosen as output measurement.

---

[12] These measures of labour productivity should be treated with care, because changes in output can also be affected by changes in inputs other than labour (e.g. capital). Furthermore, the extent to which the capacity of inputs is used can affect labour productivity; for example, there will be an apparent increase in productivity when an input that was previously not fully used becomes fully used.

In spite of this restricted definition, and in a framework that impose a strong budget constraint, the proposed productivity indicator could be useful if the aim of General Government is to deliver the same quality and quantity of services at a lower cost.

6. Concluding remarks

A methodology has been proposed to construct input and output measurement for the Agency. The analytic tools are derived from national account methodologies and are standardized at the international level. We have adopted a restricted definition of output related only to the efficiency in providing services to the users. In the future we will try to enlarge this definition in order to capture the effectiveness (outcome) that consists in providing the right output well (e. g. the dynamic of tax compliance).

Further improvements are needed in the procedure of calculation of output and input. In particular, it is necessary to integrate the gross output estimates with a measure of the net output (or value added) and to use a more general model to construct the input of labour (translog function). Furthermore, in order to conduct analysis on a longer period, the fixed weighting scheme should be replaced by chained indexes.

Despite these limits, the proposed measures can be used in a principal-agent model, and they let the reduction of the multi-task to a single-task problem. The proposed improvements allow the activities aggregation in a single contract, without facing the re-allocation of effort among tasks with different measurability degree. Furthermore and symmetrically, it reduces the agent's expected reward variance. In particular, the calculation of the production at constant price is equivalent to impose a rule of proportionality between the incentives offered to achieve each task. A fixed developing proportion between tasks implies that:
1. The performance indicator is congruent with Agency's strategy;
2. The aggregation process lowers the noisiness of the indicators.

As mentioned in the 3 paragraph, this measure does not ensure the achievement of the first best optimum, because such condition does not depend only on measurement of output but as well on: agent's risk aversion, high error term variance, task substitutability and non contractible welfare function.

The production at constant price methodology allows us:
1. To restrict informative problems (moral hazard);
2. To leave the budget based incentive system;
3. To promote the international comparability among the fiscal authorities (using common standards);
4. To realize comparison with similar market services (e. g. financial intermediation and other business activities).

Results listed in table 1 are consistent with the representation in figure 4. The new equilibrium (point B) is at higher level of production (x – expressed through the rate of growth of output at constant price) and reward (W). The reward rising is due to a reduction in labour input (THAWCP), given a constant financial endowment. Point B lies on the same Agent's reservation utility level and on a higher principal's indifference curve. This result has been caused by a positive productive shock: the organizational change connected with the creation of the Agency.

A contract based on a single output indicator could reduce the risk faced by the Agent, lowering the stochastic term variance of the production function x (see par. 3), and in this way it allows the principal to reduce the risk premium component of the reward. The realised resources could be spent both to increment the output and to reduce the financial endowment.

The reliability degree of the output measurement plays a crucial rule in the "principal-agent" model. The proposed methodologies allow us to perform an analysis of the quality process to construct the output estimates. To realize this analysis, the steps must be as follows:
1. Evaluation of the subjectivity's degree of volume indicators;
2. Analysis of the existing link between the volume indicator and the target that can be achieved;
3. Screening of the procedure accuracy to collect the basic data necessary for the construction of the volume indicators;
4. Check of the reliability degree in the estimation of the " average time of work";
5. Analysis of the linkness between the adopted weighting scheme and the principal's preferences.

The main operational goal is to improve the data collection quality and the congruity between performance indicator and principal's objective function. The expected results concern a raise of effort intensity and an improving in effort allocation among tasks.

11

References

Baker George P. (1992), Incentive Contracts and Performance Measures, The Journal of Political Economy, Vol. 100, No 3 (Jun 1992), 598-614.

Baker George, Gibbons Robert, Murphy Kevin J. (1994), Subjective Performance Measures in Optimal Incentive Contracts, The Quarterly Journal of Economics, Vol. 109, n. 4 (Nov 1994), 1125-1156.

Dixit Avinash (1997), "Power of Incentives in Private versus Public Organizations", *The American Economic Review*, Volume 87, Issue 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association, May, (378-382)

Dixit Avinash (2002), "Incentives and Organizations in the Public Sector. An Interpretative Review.", *The Journal of Human Resources*, Fall 2002 (37,4), (696-727)

Eurostat (2001), *Handbook on price and volume measurement in national accounts,* Louxembourg

Feltham, Xie (1994), Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations, The Accounting Review, Vol. 69, n. 3 (Jul 1994), 429-453.

Griliches Zvi, (eds. 1992) *Output measurement in the service sector*, The University of Chicago Press, Chicago

Hill T. P. (1977) "On goods and services", *Review of Income and Wealth*, vol. 123, pp.315-338..

Holmstrom B., Milgrom P. (1991), Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design, Journal of Law, Economics and Organization, 7 (Special Issue), 24-52.

Holmstrom Bengt (1982), Moral Hazard in Teams, The Bell Journal of Economics, Vol. 13, n. 2 (Autumn, 1982), 324-340.

Holmstrom Bengt, Milgrom Paul (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica*, Volume 55, Issue 2, March, (303-328)

Ittner Christopher D., Larcker David F. (2002), Determinants of Performance Measures in Worker Incentive Plans, Journal of Labor Economics, Vol. 20, n. 2, pt. 2.

Jorgenson D. W., Fraumeni B. M. (1992) "The output of education sector", in Griliches Zvi, (eds. 1992) *Output measurement in the service sector*, The University of Chicago Press, Chicago

Murphy Kevin J. (2000), Performance Standards in Incentive Contracts, Journal of Accounting and Economics, Vol. 30 (3).

Murray R. (1992) "Measuring Public Sector Output: the Swedish Report", in Griliches Zvi, (eds. 1992) *Output measurement in the service sector*, The University of Chicago Press, Chicago

OECD (1998), "Measuring Output for Non-Market Services", Joint OECD/ESCAP Meeting on National Accounts, Bangkok, 4-8 May

OECD (2000), *Manual on Productivity Measurement*, OECD, Paris

OECD (2001), "Performance Measurement in Tax Administrations – Practice Note", prepared by the *OECD Committee of Fiscal Affairs Forum on Strategic Managment*, Paris

Tirole Jean (1994), "The Internal Organization of Government", *Oxford Economic Papers*, Volume 46, Issue 1, Jan., (1-29)

United Nations, Commission of the European Communities, International Monetary Fund, Organization for Economic Co-operation and Development, World Bank, *System of National Accounts* (1993), New York

Wilson James Q. (1989), *What Government Agencies Do and Why They Do It*, New York, Basic Books

12

Annex 1. Indicators used in the calculation of the Agency's output at constant price

Macro-activity	Activity	number of indicators
Tax system management		68
	Acquisition data	1
	Recording tax forms	4
	Recording inheritance tax	1
	Automated recording of tax forms	1
	Audit on rent contracts	1
	Other activities concerning the record of tax forms	1
	Property tax rebate and claims	6
	Income tax rebate and claims	15
	VAT rebate and claims	15
	Other rebate and claims	6
	Rebates allowed	1
	Automated rebates	2
	Manual rebates	1
	Rebates and claims allowed to not resident	3
	Audit on fiscal intermediaries	1
	Automated treatment of tax forms	4
	Lotteries	1
	Other executions on tax system management	1
	Fulfillments of tax collection	1
	Tax credit management	1
	Activities concerning the dismissing of service centers	1
Legal advice		24
Tax assessment		68
	Audit on the fiscal documentation	5
	Audit on taxpayers	11
	Automated partial audit	3
	Audit on public acts	2
	Audit on private acts	4
	Audit on inheritance tax	2
	Audit on judicial acts	3
	Audit on rent contracts	3
	Audit on dismissing companies acts	2
	Audit on the companies estate	1
	Audit on large companies	3
	Audit on medium companies	3
	Audit on small companies	3
	Other audit activities	2
	Audit on non profit institution	1
	Tax litigation	20
Customer service		21
	Fiscal code assignment and variation	1
	VAT assignment and variation	1
	Economic activities variation and cancellation acts	1
	Electronic transmission of fiscal forms	4
	Correcting interventions of operators	1
	Call center calls	1
	Inheritance declarations registration	1
	Private acts registration	3
	Public acts registration	2
	Judicial acts registration	1
	Registration of other kind of acts	3
	Electronic registration of rent contracts	1
	Electronic registration of notarial deed	1

13

14