

Session Number: 3
Session Title: Measurement of government output
Session Organizer(s): David Caplan, Tim Smeeding
Session Chair: David Caplan

*Paper Prepared for the 29th General Conference of
The International Association for Research in Income and Wealth*

Joensuu, Finland, August 20 – 26, 2006

**A Framework for Quality Adjustment across
UK Public Services**

Jim Ebdon and Ogho Okiti, Office for National Statistics, UK

For additional information please contact:

Jim Ebdon
1 Drummond Gate, London, UK, SW1V 2QQ
james.ebdon@ons.gov.uk
+44 20 7533 5953

This paper is posted on the following websites: <http://www.iariw.org>

1. Summary

- 1.1. This article looks at the challenges involved in adjusting measures of government output to reflect changes in quality. Examples from ONS's experiences in developing measures are used to illustrate the various methods of quality adjustment, including problems that are likely to arise.
- 1.2. The main lessons are that:
 - While producing quality adjusted measures of government output is a difficult task, it is possible to define both a reasonable methodology and an achievable work programme.
 - This work programme will start by differentiating services. Where the link between unit costs and value of a given activity is weak this differentiation may measure quality changes poorly.
 - The next steps will be to adjust further using measures based on the degree of success and/or contribution to outcomes. This will only produce representative quality adjustments if all (main) outcomes are identified and covered by the measures.
- 1.3. For adult social care, the research conducted suggests that
 - Service users are capable of estimating how much benefit they are receiving from care services and so their views should be used in preference to those of service providers or inspectors.
 - The preferences of the potential service users can be used to give relative weights for different outcomes.

Consultation - Measuring performance in our public services

The Office for National Statistics' UK Centre for the Measurement of Government Activity (UKCeMGA) is soon to embark on a widespread consultation programme, covering key areas in the measurement of performance in our public services.

We will be holding consultation exercises throughout autumn and winter 2006/07. The first will be on methodological issues common to measuring all public services. It will be followed by two further consultation exercises on measuring the output of the Education and Health Services respectively.

Further details are available at:

<http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/output/UKCeMGA.asp>

2. Background

- 2.1. The Eurostat Handbook on Price and Volume Measures in National Accounts (Eurostat, 2001) recognises, in line with the SNA, the importance of incorporating quality changes into the measure of government output. The handbook makes it clear that, among other criteria, an 'A' method needs to:
 - Be defined as detailed and homogenous as possible, especially with regard to their unit costs
 - Be quality adjusted
- 2.2. This was corroborated by the Atkinson Review (Atkinson, 2005) in the recommended principles for the measure of government output. The second principle states that;

“the output of the government sector should in principle be measured in a way that is adjusted for quality, taking account of the attributable incremental contribution of the service to outcome”.
- 2.3. There is then a consensus that quality changes need to be incorporated into the measures of government output. What is less established is how to go about this in practice. The UK's National Statistician when accepting the recommendations of the Atkinson Review emphasised the necessity of establishing an overarching framework in order to avoid *“inconsistencies in approach and in rigour between the different sectoral measures”*. As a safeguard against what he called *“the temptation of ad hoc methodologies that later turn out to be problematic”*, he called for the formation of *“a robust set of principles [...] to ensure the development of professionally rigorous methodology.”* (ONS, 2005)
- 2.4. The challenge set by the National Statistician is then, to develop this 'robust set of principles'. This paper sets out some of the developments ONS's 'UK Centre for the Measurement of Government Activity' (UKCeMGA) has undertaken with partners to measure quality changes and looks to see what this work suggests in the way of principles.

3. Measuring Quality

- 3.1. The Atkinson Review's recommendation 6.7 states that while ONS should give a high priority to work on quality adjustments, *'a relatively high threshold should be set for their introduction into the National Accounts; in particular, ONS should not introduce quality adjustments until it is assured that the dimensions covered are sufficiently representative'*.
- 3.2. Or as the National Statistician described it, *'proposed quality adjustment to output measures must be representative of the changes in the quality of a service as a whole and not a subset or a selective element of quality when its other dimensions may be moving quite differently'*.
- 3.3. The Atkinson Review offers three methods of adjusting for quality that could be used to produce a representative set of measures;
 - Differentiation of services
 - Definition of the volume measure by degree of success
 - Adjustment by contribution to outcomes
- 3.4. These methods are explored in more detail in the following sections using examples from ONS's recent articles on public sector productivity.

Differentiation of Services

- 3.5. Differentiation of services into categories that are homogenous is the most established methodology used in national accounts. This familiarity makes the method usually the easiest first step in approaching quality adjustment.
- 3.6. As the Eurostat Handbook makes clear, using output indicators which are defined as detailed and homogenous as possible with regard to their unit costs captures changes in quality arising from structural changes within the aggregate output.
- 3.7. In the UK the prime example of service differentiation is the cost-weighted activity index for health which now has over 1900 activities, up from 16 used prior to 2004. This differentiation allows the aggregated output measure to capture shifts from lower cost, lower quality treatments to higher cost, higher quality treatments.
- 3.8. The advantage of the method is that the measurement of quality changes is 'built-in' and so, providing data is available, quality changes arising from structural changes do not need to be investigated further. However, the accuracy of the method is determined by the extent to

which differences in the unit costs of treatments are representative of differences in quality. In services where unit costs are largely determined by expenditure on staff the assumption that differences in unit costs correspond to differences in quality for alternative activities may be valid. Where it is not true, that is where the alternative services are associated with high cost, low quality and low cost, high quality this differentiation can lead to misleading results.

- 3.9. In health, the recent uptake of statins (drugs used in the treatment of coronary heart disease) is a good example. These drugs are more effective than alternative treatments and the uptake of these drugs has therefore led to improved outcomes (more life-years accruing to patients) and so, health output should have increased. However, the drugs have lower unit costs than alternative treatments and so measured output has not captured this change.
- 3.10. One way to correct for this, and proposed by the Atkinson Review, is to use weights that relate to the value of the activity rather than the unit costs. Value weights are arrived at by looking at the outcomes of the activity. For Statins, while the unit cost of prescription is £27, the value added in terms of added life years is £115 (DH, 2006). The difference to measured output of using the value of statins rather than the cost is large – adding potentially 0.8 percentage points on average to the growth of health output in recent years.
- 3.11. This use of values rather than unit costs is a radical change to normal national accounts practice, as the National Statistician noted;
- ‘The proposition differs from the existing SNA guidance, which proposes, in general, the use of marginal cost weights, though admittedly with some indication that this is on the grounds of practicability rather than from arguments in principle’.*
- 3.12. In summary, differentiation of services is a powerful and practical method of capturing quality changes but where relative costs are a poor proxy for relative quality the method can lead to perverse results. The use of value weights can rectify this problem.
- 3.13. Differentiation of services can only capture structural changes and so this method cannot be used alone to fully adjust for quality. Instead it must be used in conjunction with others.

Degree of Success

- 3.14. The degree of success describes how well a service is doing compared to the last period. Atkinson described it as *“simple repackaging: a quality improvement is assumed to be equivalent to getting a larger package.*

Higher grade petrol gives you 10 per cent more miles to the litre, so it is equivalent to 1.1 units of the lower grade”.

- 3.15. For most services there will be a number of aspects of quality, relating to various outcomes attributes, which need to be incorporated and so we will look to find a package of quality indicators. The acceptability of this package will be judged by the extent to which:
- The indicators measure changes in output quality rather than process quality alone
 - They are representative of changes in the service as a whole
 - Quality changes are attributable to the service
 - The measures avoid double-counting
- 3.16. The measure of social security administration is based on the number of different types of benefits processed. After improving the measure in terms of service differentiation further developments have concentrated on measures based on the degree of success. Atkinson recommended adjusting for two aspects of quality – timeliness and accuracy. The following paragraphs discuss how well these measures compare to the above criteria.
- 3.17. The quality aspects of social security that relate to the service received by individuals can largely be described in terms of convenience. It is clearly more convenient for benefits recipients to receive their payments more quickly and more accurately. So improvements in timeliness and accuracy are improvements in output quality and not merely process quality. The question of weighting remains – how much is a 1% increase in accuracy worth? While no answer is available yet for social security, the way forward is clear: we need to determine how much a change in the quality parameters changes the outcome.
- 3.18. It is unclear, whether, the two quality changes are representative of the service as a whole. The UK’s approach will be to consult service experts and other stakeholders at an appropriate time. There is, though, the possibility that an increase in timeliness could be achieved at the expense of a fall in accuracy and so it would be ill-advised to introduce the measures separately rather than as a package.
- 3.19. Timeliness, if measured by the average time to process a claim, is largely under the control of social security and therefore changes in timeliness can be readily attributed to social security. In accuracy, however, the picture is mixed. Certain aspects such as staff error are clearly attributable but for others the relationship is more complex. A reduction in fraud can be achieved by improved detection and elimination procedures but also will depend on the level of fraud

attempted. This latter aspect will be influenced by social and economic movements outside the control of social security. The approach taken then may be to separate the different aspects of accuracy and treat them independently.

- 3.20. Double counting may arise by overlaps in coverage between the quality indicators or between the quality and quantity indicators. The two aspects of quality have no overlaps and therefore do not involve double-counting which could, for example, be introduced by a third measure such as overall customer satisfaction for example. However, it may be that timeliness is to some extent picked up already in the measure of the number of benefits processed. In a service with a large back-log of claims an improvement in timeliness should lead to an increase in claims processed and therefore the output measured. The proposed quality indicator therefore needs to be investigated for these possible effects.
- 3.21. The quality measures for social security administration are by no means complete. Nonetheless, the example does indicate that when further differentiation is not possible, adjustment by degree of success offers an achievable route to quality adjustment.

Contribution to Outcome

- 3.22. Differentiation of services and adjustment by degree of success have both to some extent needed to consider outcomes: whether by assuming that higher unit costs are associated with higher outcomes; or by using weights directly related to the outcome. An alternative is to use outcomes more directly, indeed the Eurostat handbook states that:

'The most appropriate way of adjusting for quality, therefore, is to investigate changes in outcome indicators'.

- 3.23. The Handbook also highlights two of the difficulties of using outcomes directly: attribution and time lag. An example from education illustrates both well.
- 3.24. One aspect of the quality of school output is how much levels of educational attainment are raised. A variety of potential methods of estimating these changes have been looked at in the UK based on examination scores. One such is the level of attainment of GCSE's – the standard examinations taken by UK pupils at age 16.
- 3.25. While these data are readily available, the issue of attribution is difficult to resolve. Examination success is influenced by a number of societal and economic factors; by changes in the examination structure; as well as by the efforts of schools themselves. It is clearly important to separate these influences, or at least establish that non-school

influences are relatively constant before examination results can be considered as reliable quality indicators.

- 3.26. Further, an examination success at age 16 is the result of 11 years of education leading to a considerable lag between changes in the quality of teaching and changes in outcome.
- 3.27. An outcome indicator such as examination success attempts to combine changes in the quality of all relevant aspects of education that contribute to the outcome (teacher quality, methods, materials etc.) into a single measure. It cannot therefore be applied alongside another quality measure that includes these aspects or there will be double-counting.
- 3.28. The use of outcome indicators can, from the education example, seem daunting. The following case study from adult social care demonstrates however, how they can be used as part of an overall measurement framework.

4. Case Study: Adult Social Care

Differentiation of services

4.1. As with all other measures of government output, the first step in improving the measure of adult social care (ASC) is to increase the differentiation of services. The UK measure now has 23 activities (see table 4.1) differentiated where data exists by:

- Type of service: Services delivered in residential care homes; in the client's own home; in the community; and the referrals and assessments services.
- Client group: Services delivered to younger adults with physical disabilities; younger adults with learning disabilities; younger adults with mental ill health; and older people.
- Type of home: nursing homes; local authority care homes; and independent care homes.

Table 4.1 Cost Weighted Activity Index for Adult Social Care

Activity	Weight in 2004 (To nearest %)
Referrals and Assessments	15
Older People	8
Younger Adults – physical disabilities	2
Younger Adults – learning disabilities	2
Younger Adults – mental ill health	3
Care Homes – Older People	33
Nursing Homes	10
Local Authority Care Homes	7
Independent Care Homes	16
Care Homes – Younger Adults with Physical Disabilities	3
Nursing Homes	1
Local Authority Care Homes	0
Independent Care Homes	2
Care Homes – Younger Adults with Learning Disabilities	12
Nursing Homes	1
Local Authority Care Homes	2
Independent Care Homes	10
Care Homes – Younger Adults with Mental Ill Health	3
Nursing Homes	0
Local Authority Care Homes	0
Independent Care Homes	2
Day Care	12
Older People	3
Younger Adults – physical disabilities	1
Younger Adults – learning disabilities	6
Younger Adults – mental ill health	1
Domiciliary Care	23
Total meals provided all sectors	1
Home care all provision	20
Total number of people receiving equipment	2

Contribution to Outcome

- 4.2. Differentiating services will have captured some quality changes but not all. The measure of output of ASC should be the amount of care received by clients; however, the activities are measured in units such as the number of care weeks (for care home services). Care weeks are a reasonable proxy providing the amount of care delivered each week is constant over time.
- 4.3. In the UK there is good reason to believe that the amount of care delivered each week has been changing. UK government policy has been to concentrate services on those with the greatest needs. The implementation of this policy has meant that on average care home residents are more reliant on government services than before. Assuming for now that these increased needs are largely being met, the average amount of care being delivered per week of care will have increased. So, the current index will be underestimating output growth.
- 4.4. The activity indicators therefore need to be adjusted to reflect the amount of care delivered. The Personal Social Services Research Unit (PSSRU) at the University of Kent was commissioned by the Department of Health to propose an adjustment methodology. This methodology is summarised below, more details can be found in PSSRU (2006).
- 4.5. PSSRU's proposed methodology combines measures of possible contribution to outcome (referred to as the Capacity for Benefit or CfB) with measures of the degree of success. The research to date concentrates on services for older people as they constitute the largest client group. To measure contributions to outcome it is first necessary to determine the outcomes that can be affected.

Dimensions of Outcome

- 4.6. PSSRU point out that there are both long and short term outcomes from ASC. The longer term outcomes arise from gains in health, knowledge and from the provision of equipment. However, the research concentrates on the areas where welfare gains are current. In order to produce a representative set of quality indicators it is first necessary to determine what the dimensions of outcome are. PSSRU identified nine dimensions of outcome as follows:
 - Personal cleanliness and comfort – the individual is personally clean and comfortable.
 - Social participation and involvement – the individual is comfortable with their level of emotional support, general social contact and level of community participation.

- Control over daily life – the individual can choose what to do and when to do it, having control over daily life and activities.
 - Meals and nutrition – the individual has a nutritious, varied and culturally appropriate diet with meals at regular, timely intervals.
 - Safety – the individual feels safe and secure, including fear of abuse, falling or other physical harm and fear of being attacked.
 - Accommodation cleanliness, order and accessibility – the environment is clean and comfortable and is easy to get around.
 - Employment and occupation – the individual is sufficiently occupied in meaningful activities whether these are formal employment, unpaid work or leisure activities.
 - Role support (as a carer or parent) – the individual is able to care for their dependant(s) as much as they wish without becoming overburdened.
 - Living in own home.
- 4.7. Seven of these dimensions relate to care delivered to the client, while the eighth recognises the importance of role support; the assistance given to a client in supporting their dependents¹. The final dimension recognises that for most people moving from their own home to an institutional setting represents a loss of welfare and so care homes should be ‘marked down’ to recognise this.
- 4.8. There is more work needed to ensure that these outcome dimensions relate to the bulk of the care provided, but when a definitive list is achieved the benefits of this approach are clear. Such an holistic approach will allow us to determine whether our proposed measures are truly representative of the whole and not merely selective aspects.

Measurement of Capacity for Benefit

- 4.9. Having established the outcome dimensions that care services affect, we can now begin to estimate the extent to which clients are relying on government services for positive outcomes. Or to rephrase the question: what would clients’ needs be in the absence of government services?
- 4.10. Two datasets are needed to establish this: levels of need along each of the dimensions and the relative weights. Both datasets involve a degree of subjectivity, so a decision needs to be taken about whose opinion is valid.

¹ While this is a valid dimension of outcome the research does not include this at present.

- 4.11. The levels of need, for example, can be sought using data from commissioners, providers, inspectors, clients or beneficiaries². In a market transaction customers determine the extent to which they benefit from a given good or service and so to parallel this we should seek the views of service clients. However, there may be doubts about the ability of people to accurately answer questions about how much and in which outcome dimension a service was assisting them. PSSRU found that, after surveying clients that: *overall the evidence suggested that for the most part the questions were eliciting what was intended and we had got a consistent picture of what was happening in practice. As a result we could have some confidence in estimating Capacity for Benefit.*
- 4.12. While for older people in ASC it appears that clients' views are valid, the choice needs to be made carefully as the effects on measures can be large. Table 4.2 shows a dataset of the levels of needs across different domains for recipients of 'meals-on-wheels' (a meals delivery service). The results are interesting in two ways. First, only 79% of recipients reported that the service assisted in the 'meals and nutrition' domain and 12% who reported that they received assistance felt that they would have no needs in the absence of the service. Second, recipients report benefits in each outcome dimension. Rather than casting doubt on the methods these findings instead point to the advantage of surveying service users to find out exactly how they are benefiting rather than relying on experts assessments of what the service should do.

Table 4.2: Domains of outcome and levels of need addressed by meals-on-wheels (PSSRU)

Outcome Dimension	Receives meals-on-wheels %	High needs %	Low needs %	No needs %
Personal comfort	3	3	0	0
Social participation	14	5	2	8
Control	35	22	13	0
Meals	79	40	27	12
Safety	6	2	3	2
Accommodation	6	3	3	0
Occupation	5	3	2	0

- 4.13. PSSRU's approach to gathering data for weights has been similar. Discrete choice experiments have been conducted to determine the

² The separation of clients and beneficiaries recognises that in many cases government is substituting for care services that would otherwise be provided by family or friends. In effect then the benefit of the services accrues to the family or friends rather than the client.

relevant weights of different outcomes using potential service users as a sample group.

- 4.14. One effect of the methodology chosen (to measure potential rather than actual benefits) is that the researchers have not been able to produce marginal CfB's for individual non-residential services. Instead a CfB for all non-residential aspects has been calculated including all aspects listed as day care and domiciliary care in table 4.1 and referred to as the 'home care CfB'. Instead PSSRU find different CfB's by intensity of home care (see table 4.3), so it may be possible to differentiate the services along this dimension instead of using type of service.
- 4.15. Table 4.3 shows an example of the type of results obtained by PSSRU for home care. The CfB has a maximum possible value here off 7. Table 4.4 shows similar results for care homes. As the data is very provisional the values obtained should be treated with caution. The results do though correspond to expectations: recipients of more intensive care home services have higher needs as measured by their CfB; and the needs of care home residents have increased over time.

Table 4.3: Mean estimated Capacity for Benefit of home care packages (PSSRU)

Home care hours per week	Capacity for Benefit
0-2	1.12
2-5	1.30
5-10	1.63
11+	2.90
All cases	1.92

Table 4.4: Capacity for Benefit among admissions to care homes (PSSRU)

	Personal care homes	Nursing homes	All care homes
1995 CfB	2.02	2.49	2.21
2005 CfB	2.37	2.75	2.49
% change	18%	11%	13%

Adjusting for Degree of Success and Client Experience

- 4.16. Multiplying the capacity for benefit of a given service with the number of person weeks of care delivered produces an estimate of the maximum possible amount of care which could be delivered. That is, it is an estimate of how much care would need to be delivered to meet all the relevant needs of care recipients. Clearly services will not actually deliver all this possible care and so it is desirable to adjust for changes in the degree to which services have been successful in meeting needs.

- 4.17. Further work is needed to estimate changes in the degree of success. The advantage of the holistic approach suggested by PSSRU is that the dimensions of success and the weights have already been established making the process more straightforward.
- 4.18. Additionally there are aspects of quality that are additional to the amount of care delivered which can be summarised as client experience³. This type of quality change is best picked up in satisfaction surveys which can capture valued changes in aspects such as: the environment in which services are delivered; whether services are user-friendly; whether care workers are respectful and understanding etc.
- 4.19. In the UK data that are fit for purpose are not available for these adjustments. Satisfaction survey data for home care exist but may combine elements of care delivered and client experience whereas it would be preferable to have these separated. For care homes there are data on standards and work is underway to see if these can provide the type of quality data needed.

³ DH 2006 has an analogous division of quality changes into health effects and patient experience

5. References

Atkinson (2005) *Atkinson Review: Final Report Measurement of Government Output and Productivity for the National Accounts*

DH (2006) *Healthcare Output and Productivity: Accounting for Quality Change*

Eurostat (2001) *Handbook on price and volume measures in national accounts*

ONS (2005) *Letters and official response to the Atkinson Review.*

http://www.statistics.gov.uk/about/other_letters/default.asp

PSSRU (2006) Netten, A.; Forder, J.; and Sharpiro, J; *Measuring PSS Outputs for Adults.*