

The full cycle of a performance-measurement system: the ‘Use of Resources’ assessment for English local authorities

David Heald (University of Aberdeen) and Ron Hodges (University of Sheffield)

Paper presented at the 13th Biennial CIGAR Conference, 9-10 June 2011, at the University of Ghent, Belgium

**Contact author: Professor Ron Hodges, University of Sheffield Management School, 9 Mappin Street, Sheffield S1 4DT, United Kingdom.
Email: ron.hodges@sheffield.ac.uk**

1 INTRODUCTION

Performance measurement systems are recognised as one of the most important features of New Public Management (NPM). In relation to English local authorities, the Use of Resources (UoR)¹ was from 2002 until 2010 a comprehensive and ambitious assessment of that subset of NPM which is labelled New Public Financial Management (NPFM). The UoR was nested within broader performance assessment systems, the Comprehensive Performance Assessment (CPA) for local authorities and the later Comprehensive Area Assessment (CAA) for public authorities in given geographical areas, including local authorities.

Over a period of approximately 10 years, it is possible to track the beginning, evolution and then abrupt termination of the UoR (and its CAA parent) by the incoming Conservative-Liberal Democrat Coalition Government. From the moment of the announcement of abolition in May 2010, the CAA and UoR were immediately stopped in mid-process; auditors did not report their assessments for 2010. In August 2010, this decision was followed by the announcement of the impending abolition of the Audit Commission (AC), the central government ‘agency’² which had managed these assessment processes on behalf of central government. There had been a UK Labour Government from 1997 to 2010, meaning that these assessment systems were closely associated with that Government.

There are two important points of context. First, since 1999 the United Kingdom has had Devolved Administrations in Scotland, Wales and Northern Ireland, responsible for most public expenditure in their jurisdictions other than social security and defence, whereas England remains the direct responsibility of the UK Government. In relation to England, there has been much more governmental enthusiasm for centralised assessment regimes and for using market-type mechanisms in public services than in the Devolved Administrations. The control, supervision and financing of local authorities is devolved to each Devolved Administration. Second, there is often an assumption that intensive performance assessment systems are used for purposes of cost reduction. However, the life span of the UoR coincided with a period of remarkably rapid growth in local government expenditure in England; in this period, the Treasury’s UK public expenditure control aggregate (Total

¹ This paper uses the terminology of ‘Use of Resources’ throughout, though the Audit Commission only used it from 2005 when the ‘harder test’ Comprehensive Performance Assessment was established.

² ‘Agency’ is used here in a non-technical sense to indicate a public body at some distance from the ministerial department; the Audit Commission is a public corporation on UK classifications.

Managed Expenditure) rose by 72% in nominal terms and by 39% in real terms between 2001-02 and 2009-10 (Treasury, 2010, Table 4.1).

The focus of this paper is what went right and what went wrong with the UoR system of assessing NPFM-type aspects of local authority performance. There is inherent interest in such an important episode in a country with significant NPM experience. Moreover, this paper focuses attention on understanding the relevance of this English experience to other jurisdictions, and its relevance to whatever – if anything – might follow in England. It uses public domain information about the UoR in England as a whole. It is also informed by extensive fieldwork in one English region (Abu Hasan et al, 2010) and by the experiences of the authors as minor participants in this evolving narrative.³

This paper is structured in the following way. This Section explains the motivation of the paper and how it is structured. Section 2 considers why a central government might operate a performance measurement system for local authorities, and how this applies to England. Section 3 describes the key elements of the UoR system, and then presents the results of the UoR. Section 4 interprets the full cycle of the UoR, focusing on strengths and weaknesses. Given its abolition, there is obvious interest in the extent to which the UoR contributed to the abrupt termination of the CAA during the 2010 assessment. Section 5 considers what lessons might be learned, beyond the obvious point that systems associated with one Government may be vulnerable to a change of central government. It also considers how the analysis might be further developed.

2 CENTRAL GOVERNMENT MEASUREMENT OF LOCAL AUTHORITIES' PERFORMANCE

The generic issues concerning the measurement of public sector performance are not the principal concern of this Section. Those include: the fact that resources are scarce in relation to needs and aspirations; the view that the public sector lacks the inbuilt efficiency-promoting incentives that characterise a competitive private sector; and that public sector accountability mechanisms are operationally weak.

The key issue here is why central government would establish a system for measuring the performance of democratically-elected local authorities, rather than leaving this to the

³ David Heald has been a member since 2003 of the Audit Commission's Technical Advisory Group, which discussed plans and processes for successive UoR assessments. Ron Hodges is currently specialist adviser to the Communities and Local Government Committee of the UK House of Commons, which has mounted an inquiry into the future of local authority audit and inspection after the abolition of the Audit Commission. The views expressed are entirely their own.

accountability mechanisms applicable to local authorities. Having established legally separate and democratically elected local authorities, why would central government embark on an ambitious programme of assessment of their resource use and service delivery? On a 'localist' view, elected local authorities would take their own decisions on tax rates and on the level and mix of expenditure.

The following reasons why central government may wish to operate a performance measurement system are expressed in terms specific to England. However, some of these arguments may apply in other countries where there is multi-level governance.

First, devolution notwithstanding, the United Kingdom has a highly centralist political tradition in which the mandate of holding power at the UK Parliament is much more important than local mandates. The sovereign UK Parliament sets the statutory framework for English local authorities and can abolish/merge authorities, remove functions, reduce grants and change electoral systems without local consent. In conflicts between central and local government, which have been extensive and bitter, particularly in the 1980s, central government will eventually win because of superior legitimacy and instruments of power.

Second, much of the money spent by English local authorities comes from central government in the form of grant. Although block grants dominate, their existence creates an expectation in central government that there will be "something for something".⁴ The council tax (residential property tax levied by local authorities) is extremely unpopular politically. This attention is disproportionate to its yield, perhaps because of its visibility, leading to electoral promises to freeze council tax across the country. The council tax base has not been revalued since its introduction as replacement for the community charge⁵ in 1993, and there have been various schemes for capping and freezing. Moreover, there is a strong tradition, embedded in legislation, of resources and needs equalisation between local authorities, delivered through central government grant, in order to remove inequalities between jurisdictions.

⁴ Of revenue expenditure (£98.12 billion) by English local authorities in 2008-09, only 25% was financed by council tax (gross of council tax benefit paid to low income households). In contrast, 54% was financed by central government grants and 21% through redistributed non-domestic rates (ie non-residential property tax) pooled across England (ie local authorities receive an equalised per capita amount, not the proceeds in their jurisdiction) (Department of Communities and Local Government, 2010, Table 3.2e).

⁵ The community charge, popularly known as the poll tax, had earlier replaced the domestic rates (levied on notional residential rental values). The council tax is levied according to a schedule of bands of capital value.

Third, notwithstanding substantial removal of functions through privatisation or transfer to appointed bodies, local authorities in England still have a wide range of functions. This makes them important deliverers of core welfare state and infrastructure services. Compared with local authorities in many countries, English local authorities have large populations. The centrality of the services they provide generates concern about service performance and financial health from outside the boundaries of that jurisdiction. In such a context, there are likely to be limits on local choice if there are (a) significant geographical spillovers (eg poor schools in one authority may create population pressure/transfer pressure on neighbouring authorities) or (b) effects on the growth potential of that jurisdiction (via the labour market) or costs to the central government social security budget (via unemployment and social deprivation). Moreover, international peer review and league tables (eg OECD on school performance) aggregate data for a country as a whole; again it is central government that is perceived to be responsible. The OECD also makes comparisons of efficiency and effectiveness in public sector health provision, but healthcare delivery in England is not a local government function.

The combination of concerns about equity and about labour market skills creates a strong presumption within central government that it must monitor local authority performance. England in the 21st century has a more integrated economy and more population mobility; those educated in one jurisdiction do not necessarily work there (generated human capital has legs) and those who vote for policies with damaging consequences (eg incompetence or extravagance leading to high debts) may not face the consequences. Put more formally, interjurisdictional spillovers of costs and benefits have become more important, meaning that it is unlikely that central government could sustain a lack of interest in local authority service performance and resource use.

Fourth, it can be argued that local elections in England do not function as accountability mechanisms in relation to performance, even if they do confer some legitimacy on local authorities. Turnout is very much lower than in UK Parliamentary elections. Moreover, voting is mostly determined by 'Westminster' issues (local elections are effectively nationalised), so that councillors are not rewarded for good resource use (competence) or punished for bad (incompetence). There is considerable rhetoric about local mandates, but such claims have little impact in cases of serious conflict with central government.

Fifth, a specific factor in 2000 was that the Labour Government believed that public services, including local authority services, had been deprived of resources in the previous 20 years, with damaging consequences. It wished to increase rapidly public expenditure on health and education, and with those central resources would come more systematic monitoring. However, neither the UoR nor its CPA/CAA parents drove the distribution of public money, which differentiates them from the Research Assessment Exercise (RAE) which has assessed research quality in UK universities since 1986.

3 KEY ELEMENTS OF USE OF RESOURCES

This Section describes the structure and key elements of the UoR system, including the structural changes in its composition during the operation of CPA and CAA. It then provides a summary of the UoR results for the different types⁶ of local authorities in England. The intention is to provide a description of the movement in UoR scores during the operation of the system. This saw scores move upwards, generally, but with sharp reversals in 2005 and in 2009 when the scoring criteria were changed. The data analysed here are the overall UoR scores. The theme scores which are used to construct the UoR result are also in the public domain on the AC's web-site. However, the Key Lines of Enquiry (KLoE) scores which determine the theme scores are not available, except for 2009. It is not the intention here to report the scores of any individual local authority. An analysis of UoR results in one region of England, together with the perceptions of the local authority auditees about the UoR system, is provided in Abu Hasan et al (2010).

The UoR is the component of the CPA, and later the CAA, which assessed how well a local authority manages its finances and provides value for money (VFM). The assessment of UoR was conducted on an annual basis by auditors appointed by the AC. The UoR system was applied from 2002 to 2004 to upper-tier local authorities (Audit Commission, 2002), described here as Single-Tier and County Councils (STTC). From 2005 to 2008 the UoR system was extended to include lower-tier local authorities, described here as District Councils (DC), as part of a more challenging inspection regime called 'CPA – The Harder Test' (Audit Commission, 2005a). In 2009 each local authority received a UoR score based on a revised approach (Audit Commission, 2009b) which formed part of an organisational

⁶ In parts of England there is one tier of local government (Single-Tier) but in other parts, mostly more rural areas, there are two tiers (the County Council being the upper tier, the District Councils being the lower tier). There are several District Councils within each County Council, having smaller populations and less extensive functions.

assessment within the broader CAA system (Audit Commission et al, 2009). The abolition of CAA in 2010 saw the end of the UoR assessment system. In summary, the UoR assessment formed part of the CPA, initially for STTC and later for DC as well, then became part of the CAA covering all local authorities.

The UoR assessment was carried out by auditors chosen by the AC, which might appoint its in-house auditors (still widely known as ‘District Audit’, though the label is officially redundant) or a private sector firm. The UoR from 2002 to 2004 was divided into five themes, covering financial standing, systems of internal control, standards of financial conduct, financial statements, and legality of significant financial transactions. Each theme was sub-divided into areas of evidence; the auditors judged each area to arrive at a theme’s score of 1 (weakest) to 4 (strongest), with the theme scores then used to determine the overall UoR score, also from 1 to 4.

The way in which the UoR score was assessed was altered under ‘CPA – the Harder Test’. Five themes were retained in the revised structure, although all of the original themes were enhanced and renamed as Financial Reporting, Financial Management, Financial Standing, Internal Control and VFM. The auditor judgements to arrive at the theme scores were restated as KLoEs, with every theme having several KLoEs. The evidence for each KLoE was then elaborated for each level of score to be awarded; these elaborations were called ‘descriptors’ for the VFM theme and ‘criteria’ for the other themes within the UoR (Audit Commission, 2007). The assessment of each theme was based upon a four-star categorisation from 1-star (inadequate performance), 2-star (adequate performance), 3-star (performing well) and 4-star (performing strongly). The five theme scores were then combined to determine the overall UoR score from one to four. For example, an overall UoR score of 1-star was applied if a local authority failed to achieve at least a 2-star score on at least three themes, while to gain a 4-star UoR score a local authority had to score at least two themes at 4 stars and have no theme below 3 stars.

The introduction of CAA in 2009 resulted in a further revision of the structure of the UoR assessment (Audit Commission, 2009b). The UoR was split into three new themes on Managing Finances, Governing the Business and Managing Resources. The changes reflected a broadening of the meaning of use of resources in this context. For example, the Governance theme included commissioning and procurement procedures that would deliver

sustainable outcomes and value for money, while the Managing Resources theme included effective use of natural resources to cover each authority's impact on the environment, the management of physical assets and workforce development: "From 2008/09, the use of resources assessment is more demanding than previous assessments" (ibid, p.6).

Table 1 provides a summary of the UoR scores achieved by local authorities, the first section for STTC and the second for DC. From 2002 to 2004 the CPA system applied only to single tier and county councils. The national results in that period show an increase in mean scores from 3.02 to 3.67. When 'CPA – the Harder Test' was introduced in 2005 there was a reduction in the national mean score to 2.57, followed by an increase over the next three years to a mean of 3.13. The new form of UoR, as part of CAA, saw the mean score reduce to 2.55 in 2009. The second part of Table 1 lists the scores for DC, which were UoR-assessed for the first time in 2005 with a mean national score of 2.32. The mean UoR scores increased to 2.70 over the next three years before falling to 2.27 on the introduction of the revised form of UoR under CAA in 2009.

Table 2 reflects the impact of changing UoR scores on individual councils by indicating the number of local authorities which received increased, constant or decreased UoR scores in subsequent years. Unsurprisingly, this table shows that most years saw the majority of councils scoring the same or an improved UoR score. There are two notable exceptions. In 2005, with the introduction of 'CPA – the Harder Test', 133 out of 150 STCC authorities saw their score reduce. In 2009, with the broader UoR test under CAA, 87 out of 147 STCC authorities and 102 out of 201 DC authorities saw a decline in their overall UoR score.

4 INTERPRETATION OF THE FULL CYCLE OF USE OF RESOURCES

Reviewing the life span of the UoR, how are these events to be interpreted in the light of the then UK Government's purposes and the practical experience of running such a comprehensive performance assessment system for local authorities? Some of the issues are likely to be generic, whereas others will be highly specific to the English system.

First, the fundamental problem is knowing what constitutes success and what constitutes failure, and this issue permeates all others. Success measured against low expectations might be worse in absolute terms than failure measured against high expectations (Hofstede, 1968; Otley, 1987). This means that a nuanced view must be taken of results against

performance criteria. However, the politics of (apparent) failure are difficult to handle, especially for elected politicians.

It is difficult to avoid the question whether increasing scores do represent genuine improvements in underlying performance. They might be entirely due to grading drift, resulting from (a) learning by auditees of how the scoring system works, and/or (b) ‘undesirable’ behavioural modification, and/or (c) fraudulent manipulation of the performance numbers. Moreover, there may be features of the performance measurement system which falsely create the impression of deteriorating performance, as when the UoR scores dropped sharply on the transitions from ‘original’ CPA to ‘harder-test’ CPA, and then on the transition to CAA.

Those running the performance system and those subjected to it are intelligent actors who learn how the system works, meaning that game-playing operates on both sides. For example, the AC had struggled with its policy that all local authorities should have audit committees, and achieved this objective by making their existence contribute to UoR scoring. Abu Hasan et al (2010) found that local authority finance directors thought carefully about the practicalities of improving their scores. Hood (2007) suggested that those designing a performance assessment system should keep some uncertainty as to exactly what will be measured, to offset the tendency to focus on the measured and to neglect the unmeasured. Inevitably, a rounded evaluation of whether a particular performance assessment system has brought genuine improvement has to look for supporting data, not just at the reported scores.

The second issue relates to the attribution problem, which plagues attempts to evaluate the success of individual instruments. This is particularly important in the case of the UoR, given its nesting within, first, CPA and then CAA. The public profile of UoR itself was low, seen by those outside the public financial management community as merely a building block for CPA/CAA scoring. To illustrate the issues, let us *assume* that local authority performance did improve in the period 2001 to 2010. Even on this basis, it is difficult to confirm a causal chain that runs from UoR to that improved performance. It might be argued either that this should be attributed to the higher-level CPA/CAA or to the large increase in local authority spending over that period.

There is no conclusive answer to the attribution problem when there are multiple changes over the life span. All that can be done is to examine the plausibility of various narratives. In this specific case, what happens to ‘underlying’ performance from 2010 onwards will not be conclusive. Even if there were, for example, a marked deterioration in the timeliness and quality of local authority accounts and a number of sudden financial collapses, it may not be clear to what extent this stemmed from the abolition of the UoR, or the abolition of the CAA, or the abolition of the AC, or from severe reductions in central government grants to local authorities.

Third, attention must be paid to the substantive content of the measurement instrument. A difficult-to-answer question is how well this is aligned with the requirements for effective resource use and service delivery, especially how these are perceived by auditees. To answer that question definitively requires understanding of, and consensus on, those requirements. Where that exists, there is limited scope for dysfunctional consequences arising from chasing the indicators rather than improving service performance. A contributory factor will be the extent of alignment of the measurement instrument with the professional standards, norms or expectations articulated by relevant professional bodies. Abu Hasan et al (2010) reported that some of their fieldwork respondents related UoR requirements to best practice as promulgated by the Chartered Institute of Public Finance & Accountancy (CIPFA), the professional body to which many local government finance professionals belong. There was also reference to the Audit Commission’s (2005b) own publication on *World Class Financial Management*.

As long as the measurement instrument stays on ground where there is at least some professional consensus, the potential problems associated with central government assessment of local authority performance may remain subdued. In the case of the UoR and its parents, this performance monitoring had been contemporaneous with a rapid increase in resources. However, the UoR and its parents were also driven by the political objectives and priorities of the incumbent Labour Government. Paradoxically, a performance monitoring system that is judged successful may result in expanded ambitions which appear to auditees to be instruments of policy control. An important characteristic of English local government is the extent to which political parties in opposition in the UK Parliament win control of an increasing number of local authorities. This occurred during the 1979-97 Conservative Government and again during the 1997-2010 Labour Government. By 2010, Labour had

lost most of its councils and the Conservative and Liberal Democrat parties were dominant, including having political control of the Local Government Association (LGA). This sharply altered the context from that in 2000. Whereas the UoR and CAA may still have enjoyed the broad support of local government professionals, there were partisanship reasons to expect greater resistance to central government ‘intrusion’. The introduction of the CAA in 2009, with the marked reduction of scores on the UoR and on CAA in relation to CPA, was unfortunate timing in relation to the political cycle. This contrasted sharply with the introduction of ‘CPA – The Harder Test’ in 2005, the year in which the Labour Government won its third consecutive term.

Fourth, there is the question of which organisation runs the performance measurement system, and in what style. This is a generic question but has particular salience to the UoR, which involves one tier of government assessing a ‘subordinate’ tier that claims its own democratic legitimacy. Organisationally, the assessor could be: the central government itself through a department; an ‘agency’ of central government; an audit body responsible to the Westminster Parliament or one that has judicial status; or by local authorities themselves (eg through the LGA, which is a membership organisation to which most authorities presently belong). In terms of style, this could, *inter alia*, be authoritarian, consultative or consensual. Brute power can be derived from legislation and from ministerial commitment. However, legitimacy is crucial if brute power is not available on a continuing basis. This requires the ability to persuade and to gain at least reluctant consent, and to sustain the continuing credibility with auditees of those who make the assessments. In practical terms, this means holding on-side, or at least neutralising, key stakeholders. In the UoR context, those stakeholders include local authorities themselves, their representative bodies, professional associations and key figures in political parties.

An interesting comparison can be made between the UoR and the RAE for UK universities.⁷ The credibility of RAE assessments has been maintained through the judgements being made by academic peers, whose competence is generally acknowledged, and who thereby enjoy respect, authority and confidence from (most of) those assessed. The UoR context is more complicated because the AC conferred the assessment role on auditors, some its own

⁷ The RAE has now been replaced by the Research Excellence Framework (REF), the first of which will cover the period 2008 to 2013. There were RAEs in 1986, 1989, 1992, 1996, 2001 and 2008. This predated the local authority performance assessment systems and, in marked contrast with the UoR and CPA/CAA, is used for distributing research funding on a non-linear scale that is explicitly designed to reward higher RAE scores.

employees acting independently, and others from private audit firms. Moreover, the UoR stretched well beyond what auditors do for financial certification audit. Whoever they are now employed by, these auditors generally have extensive public sector experience and may be accepted by local authority finance professionals as their peers, provided that the assessment criteria conform to professional (eg CIPFA) conceptions of best practice. A significant difference from academia is that the political leaderships of local authorities are entirely outside this professional community; this inclined them to self-congratulation when scores were high and to externally-directed hostility when scores fell. The alienation of key stakeholders was evident when the CAA and UoR were abolished and when the forthcoming abolition of the AC was announced.

Fifth, there are conceptual issues in relation to measurement. Power (2004) has written insightfully about the role that measurement plays in modern economies and societies. The issues that he raises resonate in relation to the UoR. Measurement is fundamental to economic and social functioning but can become mesmerising, particularly when extended to activities that do not satisfy necessary conditions for objective measurement. The results become detached from the underlying activity and the focus of auditors and auditees turns to the parallel world which the system has constructed. Measurement systems can therefore lead to (a) effort displacement (chasing the indicators rather than managing the underlying activity) and (b) outright falsification that goes well beyond what one might view as game-playing.

Whether the metrics are accepted as 'evidenced-based' and 'fair' may have a significant impact on the net effect (ie beneficial minus dysfunctional) of a performance measurement system. Whether 'naming or shaming' will improve or further damage performance depends on context, including which feasible actions are available to poor performers. Measurement systems may contribute to changes in the distribution of power within organisations: there is likely to be more centralisation of authority; managerial changes in poor performers; and labour market power conferred on those seen as high performers. Moreover, a constant emphasis on measurement may be interpreted as signalling a substitution of transparency (one of the claimed objectives of measurement) for trust, thus encouraging behaviour to become more instrumental. Measurement systems make explicit – perhaps even newly

create – winners and losers. This raises an important question of whether those that can never win can exit from the game.⁸

Sixth, there is the question of the expected durability of such performance measurement systems. An observed characteristic is one of frequent change, sometimes in specifics but at others in fundamental ways. Hood (1994) suggested that control systems may have tendencies to auto-destruct, countering the view that learning will enable those running them to introduce improvements on the basis of experience. The UoR exhibited these tendencies, though whether they were necessarily fatal is more difficult to assess.

Hood's auto-destruction hypothesis encourages researchers not to expect continuity but to look for change, and to map cycles of performance management if these exist. There is an observable pattern of constant change, characterised by promoters as continuous improvement or innovation, and couched in persuasive language that makes it difficult to resist without exposing oneself to accusations of negativity. Focusing on the UoR rather than on the higher-level CPA/CAA, certain features that might reasonably be associated with auto-destruction can readily be identified:

- As an annual system, for which there are eight years of reported data, there are suggestions of grading drift. For example, for STCC, UoR mean scores rose monotonically from 3.02 (2002) to 3.67 (2004) under the original CPA and then from 2.57 (2005) to 3.13 (2008) under the harder-test CPA. Over a short period, there developed in each case bunching *towards* the top end of the scale. Accusations of grading drift are damaging to a performance measurement system, and bunching means that the system's power of discrimination has become limited. It becomes difficult to demonstrate to a wider audience that claimed improvements are genuine
- The UoR was revamped in 2005; that year's mean score for STCC dropped from 3.67 (2004) to 2.57 (2005). The UoR mean score under CAA (2.55 in 2009) was markedly lower than the 3.13 reported for 2008. Such marked reductions when the system changes

⁸ The English Premier League, which only about six football teams could ever win, requires no-hopers to be motivated to stay in the game, notwithstanding inevitable performance failure and sometimes financial catastrophe. Motivators for no-hopers include the prospect of European football and a lucrative share of television rights. Football graphically reminds us of the distinction between (a) organisational performance (a top team may do well notwithstanding a bad manager) and (b) managerial performance (keeping a poor team in the Premier League may be an outstanding achievement).

seem likely to confuse the external users of performance information and also to alienate auditees (eg even the best performers may feel that that the top score is unattainable)

- There is evidence of “mission over-reach”, perhaps attributable to a sense that the original CPA and its UoR component had been successful (Audit Commission, 2009a). The labelling of the second phase of the CPA as ‘the harder test’ indicated that the AC, and presumably its ministerial sponsor, regarded the performance of local authorities as still below the desired standard. ‘Raising the bar’ is readily understood as an objective, in line with the mantra of ‘continuous improvement’. However, the later move to CAA, which involved the assessment not of the individual authority but of the public bodies in that locality (including health, fire and police), signalled a large extension, taking the scope beyond the direct control of local authority professional officers and the political leadership. Although the UoR remained focused on the individual local authority, the scope extended beyond what would normally be understood as NPFM, to include ‘sustainability’. What ought to have been predictable, even without the changing political context described earlier, is that scores would again fall dramatically and that local authorities would not feel in control of their CAA scores because of the assessment being multi-organisational
- The AC commissioned an official history, granting access to archives and persons and conferring editorial independence on the author. If the resulting Campbell-Smith (2008) account of the development of the CPA and its 2005 re-launch as ‘the harder test’ is to be believed, the AC had developed a pattern of living near a cliff edge, both in terms of relations with ministers and of internal conflicts. The time period of this official history does not extend to the replacement of the CPA by the CAA, but over-confidence derived from perceived success may have made the Commission’s leadership insensitive to the perils involved. In May and August 2010, the Commission was pushed off the cliff by the incoming Government, notwithstanding it having been a Conservative Government creation in 1983

Having completed this interpretation of the life span of the UoR, the unavoidable question is what, if anything, will follow the abolished performance assessment systems. Unfortunately, there can presently be no definite answers. Over the Spending Review period running from 2011-12 to 2013-14, English local authorities are facing a 28% reduction in the real value of

central government grant, accompanied by grant penalties for above-norm increases in council tax. Beyond that, there are several possibilities:

- If the CAA and UoR were valueless, there will be no effects, other than substantial resource savings
- If the CAA and UoR had driven performance improvement, their removal might lead to a deterioration of performance generally and a number of local authority financial collapses. Scandals would cause reputational risk to local authorities collectively, create pressures for central government intervention, and possibly lead to a new generation of performance measurement
- More functions might be removed from local authorities, continuing and intensifying a long-term trend: schools are one function that is often mentioned in this regard. Functions might be privatised, mutualised or transferred to appointed bodies
- The UK Coalition Government, likely to be in office until 2015, might lose its present enthusiasm for “localism” – the centre withdraws – as the traditional cycle of local government becoming more and more controlled by opposition parties is repeated
- The concern for territorial equity, requiring reasonably uniform service delivery across England of public services, has been one factor behind centralised assessment systems. There is current policy discussion about reducing the extent of resources equalisation, thereby allowing ‘richer’ authorities to retain their resources, particularly from non-domestic property. Such a development would be radical in light of English traditions, and would have unpredictable consequences: this might be variously interpreted as a triumph of localism, efficiency and accountability, or as the source of inequalities and spillovers that demand a new round of central intervention

5 CONCLUSION

This paper reports research in progress on England-level analysis of the UoR over its life span of nine years, only eight of which were reported. This is a successor to completed research (Abu Hasan et al, 2010) on the UoR in one English region, which adopted the twin approach of questionnaire and follow-up interview. That research presented a generally positive yet nuanced view of the UoR on the part of finance professionals in 2008.

Our intention is to expand the statistical material in this paper’s Section 3; this activity has taken much longer than expected and is not yet complete. There has been a substantial data

assembly exercise from multiple sources, which has identified the need for more checking.⁹ In particular, we will develop analyses of conditional probabilities of score movements and, for the one year (2009) where such data are in the public domain, delve below the theme scores to the underlying KLoEs. Particular attention will be paid to different segments of the performance distribution, in order to look at ‘life histories’, the effects of aggregation techniques (which were algorithmic rather than weighted average),¹⁰ and at the reasons for few authorities achieving highest score, except in the last year (2004) of the original version of UoR.

Notwithstanding what still remains to be done, a number of conclusions emerge from this analysis of the UoR. At least some of these are relevant to exercises measuring financial management quality in other public services and across jurisdictions.

First, performance measurement systems can be fragile and suddenly blow up. Duration length does not necessarily lead to improvement through learning, but may instead lead to auto-destruction. Grading drift is evidenced by improvements in scores that appear implausible over short time horizons. As an annual system, the UoR may have been particularly vulnerable. When scores bunch *towards* the top end of the scale, scale recalibration and expanded requirements lead to a sudden drop in scores, thereby alienating auditees and puzzling users. When a system is acclaimed as a success, it may be particularly vulnerable to mission creep that becomes mission over-reach. Bringing ‘sustainability’ into the 2009 UoR moved the substance beyond what finance professionals would understand as ‘financial management quality’.

Second, such technical systems are located in a political context. This manifested itself in the 2010 abolition of the CAA (May) and of the AC (August), shortly after the Labour Government had been replaced by the Conservative-Liberal Democrat Coalition Government. The CAA had first reported in 2009, in the run-up to a General Election that the Labour Government was widely forecast to lose and when opposition parties controlled most local authorities and the representative LGA. There is an interesting question as to whether the effects of the expanded scope, particularly lower scores at a key point in the national political cycle, were anticipated either by the Labour Government or the AC. As a

⁹ An illustrative issue is that it is not always clear from the published data whether these incorporate later revisions, for example, owing to reporting errors or literals.

¹⁰ Aggregation is an issue at three levels: from KLoE to theme; from theme to UoR; and from UoR to CPA/CAA. Behavioural responses to the UoR might be affected by how the aggregation algorithms are specified.

public corporation responsible to the Department of Communities and Local Government, the AC was caught up in advancing Labour ministers' agendas, leading to the loss of support by key stakeholders. Throughout its existence, the AC had been viewed as preferable to its functions being carried out directly by a ministerial department. The fact of democratic elections differentiated this context from that of the centrally-appointed and centrally-directed National Health Service, where there were parallel assessment developments over the same time period.

Third, there is ambiguity in the accountability chain. If this is conceptualised in principal-agent terms, there are severe difficulties in determining who should be regarded as the principal (central government, electors for central government, electors for local government) in relation to the local authority as agent. Accountability can be instrumental (like a managerial mechanism), or a matter of high principle (validating the legitimacy of decision makers and their decisions, whatever these may be). This accentuates the technical problems relating to measurement and the fundamental question about what constitutes success. Performance assessments can be informed by evidence but the exercise of judgement is unavoidable, particularly in terms of weighing up different items of evidence to reach a summary score. This emphasises the importance of confidence in those making the judgements; the UoR involved professional auditors making judgements on matters far removed from mainstream audit activity. When judgements are thought to have political consequences, those judgements will be questioned and/or the overall system will be challenged.

Fourth, in the rhetoric supporting the establishment and continuation of performance assessment systems, there is frequent appeal to accountability and transparency. Securing the engagement of the wider public is exceptionally difficult and the results of CPA and CAA, both including the lower profile UoR, were detached from the local political process. The notion that higher transparency about council performance would lead to greater trust and electoral participation is unfounded (Abu Hasan et al, 2010).

Fifth, scoring systems which aim to improve transparency about the performance of public sector bodies have to live within a hostile media environment. Some of this may be ideological hostility to the public sector, some that failure is much more newsworthy than success. In the media portrayal, all units should be "better than average", otherwise there is

condemnation. Paradoxically, such systems may record genuinely improved performance (which can in part be plausibly attributed to them), yet create the impression of deteriorating performance because of their media portrayal. Formal reporting systems make a narrative of failure easy to construct. This sequence of events may make key stakeholders – here political leaderships rather than professional officers – withdraw their support, without much regard for what might follow.

References

- Abu Hasan, H., J. Frecknall Hughes, D.A. Heald and R. Hodges (2010) ‘Auditee Perceptions of External Evaluations of the Use of Resources by Local Authorities’, paper presented at the *EIASM Public Sector Conference*, Copenhagen, September.
- Audit Commission (2002) *The Final CPA Assessment Framework for Single Tier and County Councils*, London, Audit Commission.
- Audit Commission (2005a) *CPA – The Harder Test Explained*, London, Audit Commission.
- Audit Commission (2005b) *World Class Financial Management*, London, Audit Commission.
- Audit Commission (2007) *Use of Resources – Guidance for Councils – 2007 Assessment*, London, Audit Commission.
- Audit Commission (2009a) *Summing Up: A Review of Financial Management in Local Government 2005-2008*, Local government national report, London, Audit Commission.
- Audit Commission (2009b) *Use of Resources: Overall Approach and Key Lines of Enquiry*, October 2009, Audit Commission, London.
- Audit Commission, Care Quality Commission, HM Inspectorate of Constabulary, HM Inspectorate of Prisons, HM Inspectorate of Probation and Ofsted (2009), *Comprehensive Area Assessment Framework Document*, Audit Commission. London.
- Campbell-Smith, D. (2008) *Follow the Money: The Audit Commission, Public Money and the Management of Public Services, 1983-2008*, London, Allen Lane.
- Department of Communities and Local Government (2010), *Local Government Financial Statistics England: No. 20, 2010*, London, Department for Communities and Local Government.
- Hofstede, G. (1968) *The Game of Budget Control*, London, Tavistock.
- Hood, C. (1994) *Explaining Economic Policy Reversals*, Buckingham, Open University Press.
- Hood, C. (2007) ‘Public service management by numbers: Why does it vary? Where has it come from? What are the gaps and the puzzles?’, *Public Money & Management*, Vol. 27(2), pp. 95-10.
- Otley, D.T. (1987) *Accounting Control and Organizational Behaviour*, London, Heinemann.
- Power, M. (2004) ‘Counting, control and calculation: reflections on measuring and managing’, *Human Relations*, Vol. 57(6), pp. 765-783.
- Treasury (2010) *Public Expenditure Statistical Analyses 2010*, Cm 7890, London, Stationery Office.

Table 1: USE OF RESOURCES SCORES, 2002-2009

SINGLE TIER AND COUNTY COUNCILS	UNDER CPA – ORIGINAL VERSION			UNDER CPA – THE HARDER TEST				UNDER CAA
	2002	2003	2004	2005	2006	2007	2008	2009
GRADE ASSESSMENT								
4	26	82	105	3	15	27	33	3
3	102	59	41	82	99	100	103	76
2	19	9	4	62	36	22	14	67
1	2	0	0	3	0	1	0	1
TOTAL	149	150	150	150	150	150	150	147
MEAN	3.02	3.49	3.67	2.57	2.86	3.02	3.13	2.55
STANDARD DEVIATION	0.60	0.61	0.52	0.57	0.57	0.59	0.55	0.55
	Note 1							Note 2

DISTRICT COUNCILS	UNDER CPA – ORIGINAL VERSION			UNDER CPA – THE HARDER TEST				UNDER CAA
	2002	2003	2004	2005	2006	2007	2008	2009
GRADE ASSESSMENT								
4				5	7	13	13	2
3				82	132	131	149	57
2				134	93	85	68	135
1				17	6	9	8	7
TOTAL				238	238	238	238	201
MEAN				2.32	2.59	2.62	2.70	2.27
STANDARD DEVIATION				0.63	0.59	0.65	0.62	0.53
								Note 3

Sources: various Audit Commission documents.

Note 1: Excludes Walsall which was not assessed for CPA in 2002.

Note 2: Excludes Isles of Scilly which was not assessed under CAA and Bedfordshire and Cheshire which ceased to exist due to local government reorganisations. The Greater London Authority, Transport for London and the London Development Agency were assessed under CAA in 2009, but not under CPA, so are excluded from these tables.

Note 3: Excludes 37 district councils which ceased to exist due to local government reorganisations.

Table 2: CHANGES TO UoR SCORES FROM ONE YEAR TO THE NEXT

SINGLE TIER AND COUNTY COUNCILS	UNDER CPA – ORIGINAL VERSION		UNDER CPA – THE HARDER TEST				UNDER CAA
	02 to 03	03 to 04	04 to 05	05 to 06	06 to 07	07 to 08	08 to 09
CHANGES TO GRADE							
RISE BY TWO GRADES	0	0	0	0	0	0	0
RISE BY ONE GRADE	71	36	0	46	30	25	1
REMAINED THE SAME	77	106	17	102	114	116	59
FALL BY ONE GRADE	1	8	100	2	6	9	87
FALL BY TWO GRADES	0	0	33	0	0	0	0
TOTAL	149	150	150	150	150	150	147
	Note 1						Note 2

DISTRICT COUNCILS	UNDER CPA – ORIGINAL VERSION		UNDER CPA – THE HARDER TEST			UNDER CAA	
				05 to 06	06 to 07	07 to 08	08 to 09
CHANGES TO GRADE							
RISE BY TWO GRADES				1	0	0	1
RISE BY ONE GRADE				66	23	38	9
REMAINED THE SAME				168	200	181	89
FALL BY ONE GRADE				3	15	19	102
FALL BY TWO GRADES				0	0	0	0
TOTAL				238	238	238	201
							Note 3

Sources: various Audit Commission documents.

Note 1: Excludes Walsall which was not assessed for CPA in 2002.

Note 2: Excludes Isles of Scilly which was not assessed under CAA and Bedfordshire and Cheshire which ceased to exist due to local government reorganisations. The Greater London Authority, Transport for London and the London Development Agency were assessed under CAA in 2009, but not under CPA, so are excluded from these tables.

Note 3: Excludes 37 district councils which ceased to exist due to local government reorganisations.