ASSESSING PUBLIC SERVICE PERFORMANCE IN LOCAL AUTHORITIES THROUGH CPA – A RESEARCH NOTE ON DEPRIVATION

Dirk Haubrich* and Iain McLean*

The UK Government places ever-greater weight on performance assessment of local public bodies through comprehensive performance assessments (CPA). However, the CPA assessment framework has been criticised for its disregard of local factors that are beyond the control of local authorities but that affect their performance. In this article, the assessment framework is described and three different studies are appraised that have investigated the link between deprivation, as one of these external constraints, and CPA performance. Suggestions are developed as to how the analysis can be improved, by extending the choice of dependent variables and explanatory variables so as to be able to use panel data as the investigative method.

Keywords: Structure, scope, and performance of government; State and local government: Health, education, and welfare; Models with panel data; Regional government analysis – general

JEL classifications: H11; H75; C23; R50

Introduction

The UK Government places ever greater weight on performance assessment of local public bodies as a policy tool. In England, assessments are now an integral part of the policy of developing what is called ‘earned autonomy’, which links some freedoms and flexibilities with a certain level of assessed performance. Specifically, the distribution of central government grant assumes a dominant importance in a local government finance system in which Whitehall funding (currently in the region of £100bn per annum) accounts for 75 per cent of local government’s financial base. The Audit Commission (AC) publishes comprehensive performance assessments (CPA) scores for all principal local authorities in England (counties, districts and unitary authorities) that aim to describe how effectively these funds have been used.

However, ever since the first round of CPA exercises in 2002, the assessment framework has attracted various criticisms. Should deprivation be taken into account when making a judgement about local authorities’ performance in delivering public services? Does it constitute an external constraint on performance that the CPA framework does not adequately address? Is performance correlated with the desirability of areas for living and working? How well does CPA work as a way of granting earned autonomy? Are performance scores valid and reliable indicators at all? Does the CPA mechanism conflict with other government policies, such as targeting resources on the most deprived areas?

This paper evaluates the relationship between CPA scores and deprivation. It reproduces the development of the CPA framework since the Local Government Act 1999, summarises the research reporting on the framework, reviews the results of the first three CPA rounds in 2002, 2003 and 2004, outlines the Audit Commission’s plans for future rounds, and finally, in section 4, outlines possible avenues of research.

1. The years 2000/1

1.1 Developing the CPA framework

During 2000 and 2001, the Audit Commission of England developed a new performance management for councils in England, called the Comprehensive Performance Assessment (CPA). It constituted a way of assessing councils’ progress in continuously improving their services, a duty placed on them by the Local Government Act 1999. In a 2001 government White Paper (Audit Commission,
2001), CPA was first formally introduced to the public and consultations commenced to come up with the final framework.

The stated objective of the CPA framework was, and continues to be, to target support at those councils that need it most and to grant ‘freedom and flexibilities’ to the better performing councils. The incentives for the latter included:

- the elimination of ‘ring-fencing’ from most central governmental grants to the local authority, which allows the latter to spend money on those areas it (rather than the Government) deems most appropriate,

- a three year exemption from subsequent audit inspections,

- an exemption from having a ‘cap’ imposed on the authority’s planned expenditure and the level of council tax it is allowed to raise from its taxpayers, and

- the freedom not to have to submit detailed service plans to central government for approval.

The CPA framework was aimed at measuring the effectiveness of councils in terms of the way they provide the services to local people and work in partnership with other local authorities, private corporations and the voluntary sector, among others. In so doing, it focused on the leadership, the systems and processes, as well as the performance of those services. Figure 1 depicts the two assessment exercises employed:

1) **Current performance.** Much of the evidence on service performance was already publicly available and was brought together in one place for the first time in CPA. The service areas (or ‘service blocks’) assessed were the six ‘core services’: education, social care for children, social care for adults, housing, environment, and libraries and leisure facilities. Where available, performance was assessed through long-standing inspectorates and auditor judgements, such as those by Ofsted and DfES for the education service area. The Audit Commission converted their existing scoring systems into its own categories, and added those service areas that it inspected itself as well as various performance indicators. Authorities would obtain a score of between 1 and 4 (with 1 being the lowest, 4 the highest) for each of the service areas.

For the first three CPA rounds (2002 to 2004), on which this paper reports, the Audit Commission arrived at these scores by ranking all 150 unitary and top-tier councils and by assigning the score depending on where on this band an authority ends up. However, the Audit Commission decided, for CPA rounds from 2005 onwards, to use fixed thresholds to determine an authority’s score.

Irrespective of how the scores are arrived at, they were then weighted, so that education and social services had a weight of 4, housing and the environment had 2, and the remaining areas 1 each. In so doing, education and social services accounted for just over half of the authority’s service score. The scores were then combined to provide an overall assessment of a council’s current performance. These ranged from a minimum score of 15 to a maximum 60 points. The continuous data were then converted to category

<table>
<thead>
<tr>
<th>Service</th>
<th>Minimum score</th>
<th>Maximum score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Social Services</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Housing</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Environment</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Benefits</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Libraries and Leisure</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Resources</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>15</strong></td>
<td><strong>60</strong></td>
</tr>
</tbody>
</table>

Note: (a) County councils do not have responsibility for housing and benefits. In these cases, the minimum and maximum total possible scores are 12 and 48.
scores, again ranging from 1 to 4. Tables 1 and 2 give a schematic summary of this process and how the weighting affects the scores.

2) Ability to improve. A second assessment concentrated on a council’s plan to improve services in the future and itself consisted of two components:

a) A self-assessment in which councils had to answer four questions with regard to what they were trying to achieve, how they had set about delivering their priorities, what they had achieved so far, and what they planned to do next.

b) An external ‘corporate assessment’ carried out by a small team that includes an auditor, an inspector and officers and members from peer councils. A total of nine areas were assessed that way, each of which obtained a score of between 1 and 4 (with 1 being the lowest, 4 the highest). These were then weighted to produce an overall corporate assessment score that ranges from a minimum of 12 to a maximum of 48. The continuous data are then converted to categorical data, again ranging from 1 to 4 (see table 3).

The outcome of this second assessment was a report that highlighted the council’s strengths and weaknesses and, again, a score and rating for its ability to improve on a 4-point scale.

The two assessment ratings (on current performance and ability to improve) were then combined to produce the overall assessment on a 5-point scale comprising the denominations ‘excellent’, ‘good’, ‘fair’, ‘weak’ or ‘poor’. Final CPA ratings were not, however, a product of simple arithmetic. Additional minimum thresholds were applied in the calculations to account for cases where one of the council’s two ratings significantly deviated from the other. In effect, these constituted value judgements about what a good performing local authority should look like. Table 4 provides an overview of the resulting scores matrix that councils could potentially arrive at with any two given individual sub-scores.

1.2 CPA and deprivation
The Audit Commission claimed that when designing the CPA framework it was taking account of local factors such as deprivation. It recognised that deprivation and other local factors can influence the quality and selection of services that local authorities provide and stated that it was “guided by the principle that councils should be assessed, not on the circumstances they find themselves in, but on the way that they respond to those circumstances” (Audit Commission, 2002a, p. 3). Yet, the Commission’s assessment framework had not made any special arrangements to adjust the scores for deprivation, neither in its judgements, nor in its statistical calculations. In fact, it stated repeatedly that there was insufficient statistical evidence to show that level of deprivation had a significant effect on authorities’ performance on any of the performance indicators (PIs) assessed (Audit Commission, 2002c, p. 2).

Even so, many local government stakeholders and respondents to the initial consultation paper argued that the CPA assessment framework should be adjusted for deprivation. The fact that the version eventually published and employed failed to do so spurred renewed and sustained criticism. Some councils (usually those in metropolitan areas with pockets of great poverty) were
worried about the potentially disastrous results they would attain during the assessment exercises they were scheduled to do, and trade unions complained that CPA ignored the problems of recruiting and retaining excellent staff in deprived areas, a fact “that directly affects the ability to provide and improve public services” (UNISON, 2002, p. 3).

The question, then, appeared to be whether deprivation can have an adverse effect on public service performance, for example because it places additional pressures on local authorities’ resources to be able to provide services in certain service areas. However, while local contexts of service output are undoubtedly significant, the focus of analysis could also be placed on the input side of the value chain. It is then possible to arrive at the opposite conclusion, because local deprivation could be said already to be taken care of adequately by central government’s adjusting the calculation of the Standard Spending Assessment (since 2003 slightly redefined and also renamed Formula Spending Share – FSS) according to authorities’ levels of deprivation. With poor authorities receiving greater financial support, it may be assumed, an equal playing field is already established before the CPA assessment even begins.

Yet, this debate remained unresolved by the time the first CPA round was carried out, and the CPA framework discussed above was left largely intact.

2. The years 2002/3

2.1 The first CPA round in 2002

The first CPAs were carried out in 2002 for all 150 single- and upper-tier councils in England with responsibility for the provision of education and social services – that is, the 33 London boroughs, the 36 Metropolitan districts, 46 unitary councils and 34 county councils. The results of these assessments were published in December 2002 (Audit Commission, 2002a; 2002b), and just over half of the authorities (76 out of 150) ended up in the top two categories.

Examining the resulting evidence, the Audit Commission stated (again) its position that while local levels of poverty, unemployment or ill health influence the quality, type and volume of services that councils and their partners provide, “for most of the performance indicators used in the assessment framework there is no evidence of a relationship with deprivation” (2002b, p. 5). To re-affirm this conclusion, the Commission also pointed out that “some councils in the most deprived areas are ‘excellent’, and some councils in relatively affluent areas are poor and weak” (ibid).

Yet a look at the statistics and graphs produced there did not quite bear out the Commission’s confidence. While approximately 35 per cent of the most deprived quartile of the 150 councils achieved either a weak or a poor CPA score, for the most affluent quartile that figure dropped to a mere 17 per cent. What is more, the 22 councils that received an ‘excellent’, for example, were not equally spread across the different types of authorities: they included 8 (out of 34) county councils, 8 (out of 33) London borough councils, but only 4 (out of 36) metropolitan boroughs and 2 (out of 47) unitary authorities. Hence, large, prosperous and ethnically homogenous councils, such as most counties, were more likely to score the top category.

Undeniable exceptions notwithstanding, the statistical tendencies and proportions that emerged from the first round appeared to confirm suspicions held by many, but not acknowledged by the Commission, that the CPA framework does not sufficiently take account of deprivation.

<table>
<thead>
<tr>
<th>CPA result</th>
<th>Components of result</th>
<th>Type of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall rating</td>
<td>5 categories (poor to excellent) based on rating from Corporate Assessment and Core Service Assessment.</td>
<td>Ordinal</td>
</tr>
<tr>
<td>Council ability rating</td>
<td>4 categories based on score from assessment of council ability.</td>
<td>Ordinal</td>
</tr>
<tr>
<td>Council ability score</td>
<td>All councils are scored on a scale of 1 to 4 on 9 different themes. The scores for each theme were weighted so that scores ranged between 12 and 48.</td>
<td>Interval</td>
</tr>
<tr>
<td>Core service performance rating</td>
<td>4 categories based on the score from the Core Service Performance.</td>
<td>Ordinal</td>
</tr>
<tr>
<td>Core service performance score</td>
<td>All councils are scored on a scale of 1 to 4 for 7 service areas. The scores for each service area are weighted and averaged so that scores range between 1 and 4.</td>
<td>Interval</td>
</tr>
</tbody>
</table>
The Commission therefore produced a more detailed account of the statistical analyses it had conducted to arrive at the conclusions it so briefly described in its CPA report. In April 2003, it produced a 27-page study exploring the interaction of CPA scores and deprivation (Audit Commission, 2003b). As predictor and dependable variables respectively, the Commission used:

- the Average of Wards Scores at district level, which is one of six outcome measures available from the Index of Multiple Deprivation 2000. According to the Commission (p. 14), the IMD level can be presented either as an interval measure – using the average of ward scores – or as an ordinal measure, for example, by grouping the ward scores into categories based on quartiles. The Average of Ward Scores was chosen in preference over other measures because it describes the council area as a whole.

- the CPA result achieved by each council, in the form of the five different data sets in table 5.

The Commission assigned councils to one of four quartiles according to their deprivation index, performed the statistical analysis using the five dependent CPA variables above and concluded that:

- there is no evidence of a difference in the rating of council’s ability to improve according to their level of deprivation (p. 4). The Commission employed three statistical tests (Chi square, Kruskal Wallis, ANOVA) as a way to assess whether the variation in the deprivation levels for councils with different ability levels ratings are likely to occur by chance. The closer the three test values, which can range from 0 to 1, are to zero, the less likely it is that any difference in council ability according to deprivation could have been arrived at by chance. For most purposes, statistical results need to have a probability of less than 0.10 before they are treated as being suggestive of a difference. The three test results above showed 0.458, 0.111, and 0.159 respectively and therefore failed to achieve that benchmark.

- there is evidence of a difference in the core service performance ratings for councils with different levels of deprivation. The three probability tests produced 0.026, 0.002 and 001 respectively as values (p. 6). The relationship between the two data sets was therefore said to be beyond the bounds of chance. However, a value of $R^2$ of 0.078 indicated that deprivation ‘explains’ only 7.8 per cent of the variation within the core service performance score. That means that deprivation cannot explain 92.2 per cent of the variation, which is why the Commission deemed the correlation to be insignificant (p. 7).

- there is no evidence of an interaction between core service performance scores and deprivation. The Commission concluded that the interaction differs depending on which type of authority is analysed and detected a clear pattern only for London Boroughs (with $R^2$ of 0.184), while county councils, metropolitan authorities and unitary authorities showed no correlation. In the case of London, however, the Commission attributed the comparatively high correlation to factors such as transitory population or pressures of recruitment that are specific to the capital (p. 9) and concluded that there is overall very little relationship between deprivation and service performance score.

- there is limited evidence of a weak correlation between the final CPA rating and deprivation. The three probability tests of chi square, Kruskal Wallis, and ANOVA showed values of 0.360, 0.115, and 0.133 respectively and the results were therefore said to be within the bounds of chance (p. 11).

### 2.2 Reactions: the Cardiff (Andrews) study

The first analysis of the 2002 CPA round that was not conducted by any of the stakeholders involved in it was a brief 8-page study by Rhys Andrews of the University of Cardiff. The report was published in the journal Public Money and Management in January 2004, a year after the Audit Commission produced the results of the 2002 CPA round and six months after the Commission’s deprivation study. In order to explore the statistically significant impact of deprivation on service performance, Andrews used:

- the Average Ward Deprivation Score (AWDS) on the Index of Multiple Deprivation 2000 for each upper-tier local authority. The AWDS is the standard population-weighted measure of deprivation used by the Office for the Deputy Prime Minister and the Audit Commission. It is derived from 33 indicators across six domains relating to deprivation, such as levels of education, income and health. These indicators were obtained from the Office for National Statistics.

- the Best Value and Audit Commission performance indicators (PIs) from 2000 and 2001 that were used in the first CPA round to assess the service areas. These
PIs are based on common definitions of key data, which are obtained by councils for the same time period with uniform collection procedures.

Note that, unlike the Audit Commission’s work described in the previous section, Andrews did not use the scores and ratings on ‘performance’ and ‘ability to improve’ or the overall CPA rating (‘poor’ to ‘excellent’) as a dependent variable. Rather, he based his regression analysis on performance only, and in that not at the score or ratings, but on the PIs underlying those scores. In CPA 2002, over 100 of such PIs were used for all of the service areas, including primary education; secondary education; mainstream schools; children’s social services; adult social services waste collection; waste disposal; planning; housing management; housing repairs; benefits and revenues; highway maintenance; and leisure and culture. In each area, between five and fifteen different PIs were applied to measure output quality, effectiveness, or customer satisfaction. For example, for the services area of ‘secondary education’, the PIs used were surplus places; unfilled places; average GCSE score; five or more GCSE at A+ to C; exclusions; and unauthorised absences. For the service area ‘planning’ they were planning applications within eight weeks; searches within ten days; and satisfaction with the planning process.

With this finer comb, Andrews came to a conclusion different from the Audit Commission, in that

- “for over half of the PIs used in the first round of the CPA, deprivation had a statistically significant influence (at the 95 per cent level) on performance” (p.22),

- “in the majority of significant cases, performance was adversely affected by deprivation, particularly for those indicators measuring formal effectiveness and output quality” (p.22),

- “for three-quarters of the education indicators, deprivation was significantly related to poorer performance” (p.23),

- “half of the PIs for housing and waste management were adversely affected by deprivation” (p.23)

- “the proportion of variation in performance for the PIs attributable to deprivation ranged from 3 per cent to 64 per cent” (p.23)

Andrews concluded that “external environmental constraints on local authority performance must be recognised for accurate and equitable performance assessments” (p.19), and that some authorities may have been misclassified by not taking sufficient account of these factors (p.24).

2.3 Reactions: the NPI & Audit Commission study

Before Andrews’ study was published in January 2004, the Audit Commission saw a need to carry out further research into the subject matter. In late 2003, possibly because of criticism directed at its own account described in section 2.1 above, it commissioned the New Policy Institute (NPI) to conduct a more thorough analysis of the relationship between CPA scores and deprivation, both through statistical analysis and elite interviewing. The resulting recommendations were to be used to improve the CPA process for implementation for the CPA round in 2005 and beyond.

The study was published in July 2004 (NPI, 2004) and its authors, Guy Palmer and Peter Kenway, approached the issue in a different way. They claimed that the CPA framework is a mathematical model that is much more complex than simply taking a number of low level scores and adding them up or averaging them. Rather, it uses a whole variety of transformations, categorisations and other statistical techniques, which makes calculations and results less transparent than desired (p.3).

More specifically, their report argued that an analytical focus on performance indicators so as to arrive at judgements about a link between CPA scores and deprivation (as arguably Andrews had done) is misleading, because the nature and methodology of the CPA model differed substantially between the assessments of the service blocks. As table 6 depicts, not all of these were carried out by the Audit Commission itself.

The assessment approaches ranged from a largely

<table>
<thead>
<tr>
<th>Service block</th>
<th>Assessment undertaken by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits</td>
<td>Benefit Fraud Inspectorate</td>
</tr>
<tr>
<td>Education</td>
<td>Ofsted and DfES</td>
</tr>
<tr>
<td>Social services</td>
<td>Social Services Inspectorate</td>
</tr>
<tr>
<td>Environment</td>
<td>Audit Commission</td>
</tr>
<tr>
<td>Libraries &amp; Leisure</td>
<td>Audit Commission</td>
</tr>
<tr>
<td>Housing</td>
<td>Audit Commission</td>
</tr>
<tr>
<td>Resources</td>
<td>Audit Commission</td>
</tr>
</tbody>
</table>

Table 6. Service blocks and assessors
judgemental approach in Social Services, informed but not dictated by performance indicators and inspection judgements, to a largely mechanical process in the Audit Commission service blocks, where the overall scores were arithmetically determined by the performance indicator and inspection scores and the mathematical rules for their combination.

This information, combined with the fact that each service block carried a different weight, convinced Palmer and Kenway that a focus on PIs alone is overtly simplistic. In figure 2 below, which is reproduced from their study, they showed that CPA scores were mainly driven by the results of inspections and assessments, with performance indicators themselves only making up around 11 per cent of an authority’s overall score (p. 13).

The information sources depicted here display in greater detail the bottom arrow that we depicted in figure 1 (‘The CPA framework’) at the beginning of this paper. Figure 2 suggests that performance indicators are of only minor importance in the overall CPA framework, somewhat qualifying Andrews’ study, which focussed solely on the link between CPA ratings and these indicators. While PIs are popular because of their transparency and suitability for basing comparisons on them, they appear to be of only limited importance in the CPA framework overall.

The question then suggests itself of how deprivation is accounted for in the other assessment methods, such as inspectorate judgements or plan assessments. Here, the NPI study concluded from interviews held with officials in local authorities that for those inspected “it is not transparent how this is taken into account”; “how it is balanced against other measures in reaching [an] overall judgement” and that confidence in the system very much depends on the quality of the inspection team (p. 25). However, the study concluded from interviews with inspection teams that in no case was any guidance given to inspectors on how deprivation should be accounted for (p. 26). It therefore recommended that in future CPA rounds, the “onus [is] on inspectorates demonstrating that they really have taken deprivation into account” (p. 28), presumably through more rigorous training efforts.

**Figure 2. Information sources used to calculate overall CPA ratings**

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall CPA assessment</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Corporate assessment</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Service Performance</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Performance Indicators</td>
<td>11%</td>
<td>6% (i.e. half), 7% (i.e. half)</td>
</tr>
<tr>
<td>Inspections</td>
<td>18%</td>
<td>0% (i.e. half), 0% (i.e. all)</td>
</tr>
<tr>
<td>Assessments</td>
<td>17%</td>
<td>0% (i.e. all), 0% (i.e. all)</td>
</tr>
<tr>
<td>Plans</td>
<td>3%</td>
<td>0% (i.e. all), 0% (i.e. all)</td>
</tr>
<tr>
<td>Audit</td>
<td>2%</td>
<td>0% (i.e. all), 0% (i.e. all)</td>
</tr>
<tr>
<td>Education</td>
<td>13%</td>
<td>6% (i.e. half), 7% (i.e. half)</td>
</tr>
<tr>
<td>Social Services</td>
<td>13%</td>
<td>0% (i.e. all), 0% (i.e. all)</td>
</tr>
<tr>
<td>Benefits</td>
<td>3%</td>
<td>0% (i.e. all), 0% (i.e. all)</td>
</tr>
<tr>
<td>The 3 Audit Commission Areas***</td>
<td>17%</td>
<td>4% (i.e. a fifth), 11% (i.e. two-third)</td>
</tr>
<tr>
<td>Resources</td>
<td>3%</td>
<td>1% (i.e. all), 1% (i.e. all)</td>
</tr>
</tbody>
</table>

* *= used to inform the assessment; ** *= used to validate the assessment; *** = combined weightings for Environment, Housing, and Libraries and Leisure

Other than that, the NPI study came to similar conclusions about the correlation between overall CPA scores and deprivation that the Audit Commission’s own report had come to a year earlier. There was “no overall relationship between the corporate assessment scores and deprivation” but a much stronger relationship between core service performance and deprivation (p. 15). However, for the core service performance, the study also stated that “if education is excluded [. . .], then there ceases to be any discernable relationship between core performance score and level of deprivation.” (p. 16). In particular, the PIs (as opposed to the judgements) that were used in the education service block demonstrated a strong correlation with deprivation \( (r^2 = 0.35) \) (p. 45). Given that “these were mostly raw educational attainment results and attendance rates, this is not surprising (p. 45). Essentially, then, CPA scores on education were significantly correlated with social deprivation, but the scores for other services, and overall scores, were not.

Palmer and Kenway therefore recommend for future rounds that Ofsted, DfES and the Audit Commission should “jointly agree whether, in principle, educational attainment and attendance rates should be included in, or excluded from, the CPA model” (p. 35). If attendance and attainment rates were to be included, then local authorities should be told why this is so and to what extent Local Educational Authorities are able to influence these scores. They also recommended that “educational attainment could be examined separately for pupils in receipt of free school meals and for other pupils” (ibid).

During interviews, the NPI team also endeavoured to identify the way in which local authorities themselves believed that high levels of deprivation affect their jobs. The resulting insights they grouped under five major headings (p. 2):

- ‘Fractionalisation’, that is, the extent to which the local population is divided between different ethnic, religious or national groups (so that councils with higher levels of fractionalisation have more languages to cope with and may have lower levels of satisfaction of services among citizens).

- ‘Transience’ or ‘turnover’, that is, the speed at which the local population is changing (so that the administration of council services, the monitoring of their impact, as well as political participation is made more difficult).

- ‘Capacity’, that is, the ability of citizens to take an active role in public life (so that individuals in deprived areas have greater difficulty to understand their rights and entitlements).

- ‘Regeneration’, that is, an activity that absorbs more resources in authorities with high levels of derivation.

- The additional challenges when working for an authority in a deprived area.

### 2.4 Reactions: the Cardiff (Andrews, Boyne et al) study

A third study that made use of the 2002 CPA data is the study by (again) Rhys Andrews of Cardiff University, this time with colleagues George Boyne, Jennifer Law and Richard Walker. The relevant article was published in January 2005 (Andrews, Boyne et al., 2005).

The study’s aim was to develop and test a model of the impact of external constraints on local service performance, that is, the extent to which success or failure in service provision is attributable to circumstances that are beyond the control of local managers and politicians. In order to answer this question, Andrews et al. identified a set of explanatory variables that are beyond the control of local policymakers, at least in the short run (p. 643). The ten variables they used were of a political, economic, social and environmental nature and included: Quantity of Service Need; Age Diversity; Ethnic Diversity; Social Class Diversity; Discretionary Resources; Lone Parent Households; Population Change; Population; Population Density and Political Disposition.

These were then linked to authorities’ scores in terms of core service performance and ability to improve. The raw scores (between 15 and 60) were used rather than the categories 1 to 4, so that the subjective judgements used to compress the data would not taint the analysis (p. 649). The authors found that the ten constraint variables collectively explain around 35 per cent of the inter-authority differences in core service performance and 28 per cent of the differences in ability to improve. They conclude that these are “satisfactory levels of statistical explanation” (p. 650).

However, the authors were particularly interested in testing seven hypotheses relating to core service performance, and the statistics failed to support these hypotheses to an equal degree. Those hypothesised correctly included:
Higher ethnic diversity and social class diversity appear to place additional burdens on service providers and thereby result in lower performance (p. 650).

Areas with declining population and a high percentage of single parent households found it more difficult to climb the CPA ladder (p. 651).

Large authorities find it easier to achieve good CPA results (p. 651).

No difference is found if the model is tested at the level of one of the four sub-groups of authorities (County Councils, Boroughs, Metropolitan, Unitaries).

The statistical results convinced the authors not only that circumstances beyond policymakers’ control had a significant impact on core service performance, but that the latter variable had a strong positive influence on inspectors’ judgements about councils’ future ability to improve. This showed that the impact of external constraints operates directly through core service performance, which in turn affects the ability to improve and the final CPA rating. In the final paragraph, the authors then conclude that,

“the CPA process is flawed by the failure to take account of circumstances beyond the control of local policymakers. The performance effects of these constraints have distorted the rankings in the CPA league table. Some councils have been falsely lauded for operating in favourable circumstances, whereas others have been wrongly criticised for the performance effects of difficult local conditions” (p. 654)

2.5 The second CPA round in 2003

None of the studies outlined in sections 2.2 and 2.3 above had any bearing on the Audit Commission’s (2003a) second round of CPAs of all 150 single-tier and county councils, because they were not yet published. As far as we can tell, the methodology used was largely the same as the 2002 round, both with regard to deprivation and any other aspect relevant to our analysis.

As to assessment results, the authorities that were inspected showed an overall improvement in their CPA scores, with 26 councils moving up one or more CPA categories, while nine moved down. In total, 55 per cent of councils ended up in the top two categories, up from 51 per cent the year before.

3. The years 2004/5

3.1 The third CPA round in 2004

In 2004, the Audit Commission conducted the third CPA round and published the results in early 2005 (Audit Commission, 2005a). For the exercise, the results of the Corporate Assessments that had been performed in 2003 were re-used to determine the 2004 outcomes. The 2004 results therefore relied more heavily on councils’ service ratings, particularly the social services and education assessments carried out by the Social Services Inspectorate and Ofsted respectively.

In the 2004 round, the rules for jumping up a CPA category were changed slightly, so that a council had to improve by at least three points to move up a category. The Audit Commission claimed that its impact was significant despite the stricter standards: of a total of 150 councils, the number of ‘poor’ and ‘weak’ councils dropped from 34 in 2002 to 16 in 2004; nine out of ten councils that were ‘poor’ in 2003 moved up at least one category in 2004; and the number of ‘weak’ councils stood at 15, down from 18 in 2003 and 21 in 2002. At the top end of the spectrum, 101 councils achieved a rating of ‘excellent’ or ‘good’, up from 76 in the first CPA round in 2002.

3.2 The first CPA round for district councils

The parts of England that have two-tier local government contain 238 non-metropolitan/shire district councils. In 2003/4, they underwent their first assessment exercise, which was completed and reported upon in August 2005 (Audit Commission, 2005b). Only 28 councils achieved the top ‘excellent’ rating, while another 86 achieved a ‘good’ rating, leaving the remaining 52 per cent with a ‘fair’ (86 councils), ‘weak’ (29) or ‘poor’ (9) score. As far as we can tell, the methodology used for districts resembled the assessment carried out previously for unitary and top-tier authorities, at least as far as the issue of deprivation is concerned. The Commission again found only a weak relationship between deprivation and councils’ performance, and failed to identify any regional factors that might justify the different achievements.

Also, the Commission found that there was no relationship between the performance of district councils and that of the top-tier authorities and that district councils serving larger populations had only a very slight advantage in their overall performance over smaller ones. Finally, the CPA results for districts were published alongside a consultation document for a new
framework to assess districts from 2009 onwards (Audit Commission, 2005c).

3.3 Plans for the future – the 2004 consultation documents

In 2004, the Audit Commission launched a consultation document about the CPA assessment frameworks in future rounds (Audit Commission, 2004a) and reported on it within a few months (2004b). In the report, the Audit Commission announced that the new CPA regime would take into account the findings of the NPI study (2004a, p. 8; 2004b, p.14); that a greater understanding of deprivation will form part of the consideration of the local context, through a better training and briefing of inspection teams (2004a, p.8); and that the work the Commission undertakes in consultation with local authorities on the topic of deprivation was already informing the intended treatment of best value indicators and inspection methodologies. The final report on this work was announced for July 2005 (Audit Commission, 2004b, p. 14) but, as of September 2005, has not yet been published.

As to other changes that the Audit Commission announced for future rounds, the new CPA regime is said to reduce inspections by more than two-thirds, leaving only the poorest services to be inspected. Corporate Assessments, which evaluate a council’s strategic approach to council business, are to be carried out once every three years for each council. This means that councils will be able to improve their CPA category without the need for a further corporate assessment, which will now be restricted to their three-year cycle.

3.4 Plans for the future – the 2005 consultation documents

In June 2005, the Audit Commission (Audit Commission, 2005b) published a follow-up consultation document to the previous one (Audit Commission, 2004a), this time on the future of the service assessment framework. The deadline for submission to this document ended on 28 July 2005.

However, once implemented, it will likely lead to a shift in the balance of evidence used, “away from a reliance on inspection scores towards assessment frameworks based on available performance information including, but not restricted to Best Value Pls” (p. 2). The Commission will also use the updated Index of Multiple Deprivation (IMD 2004). More importantly, however, the Commission plans to introduce two sets of adjustments.

The first is an area cost adjustment, through which all cost Pls will henceforth take account of differences in labour costs and business rates paid on local authority premises between authorities. Two Pls – ‘H9: Average weekly management cost of housing’ and ‘C13: Cost per visit to a library’ – are affected by this adjustment, which uses ODPM data on area cost adjustments (ACA) and GDP deflator data available from HM Treasury (Audit Commission, 2005b, p. 74).

The second adjustment is an adjustment for deprivation through which adjustments are made to each reported performance indicator where the correlation coefficient equals 0.4 or above, or –0.4 or below (p. 75). The Commission’s approach now is as follows.

“In the previous CPA assessment framework we used a linear regression model for indicators that had a substantial correlation with deprivation. The method involved ranking residuals, and therefore an authority’s scoring was based on its performance relative to other authorities. For this CPA model it is proposed that we will uplift all authorities’ reported performance for these Pls to take account of their level of deprivation. Therefore, the higher the level of deprivation, the higher the uplift, and consequently

<table>
<thead>
<tr>
<th>Performance indicator</th>
<th>Correlation (R²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-satisfaction Pls</td>
<td></td>
</tr>
<tr>
<td>BV 66b – percentage of rent arrears of current tenants</td>
<td>0.46</td>
</tr>
<tr>
<td>Percentage of private dwellings vacant for 6 months or more</td>
<td>0.46</td>
</tr>
<tr>
<td>C12a – stock turnover</td>
<td>−0.53</td>
</tr>
<tr>
<td>Satisfaction Pls</td>
<td></td>
</tr>
<tr>
<td>BV74a – satisfaction of local authority tenants with service of local authority</td>
<td>−0.58</td>
</tr>
<tr>
<td>BV75a – satisfaction with participation for all tenants</td>
<td>−0.5</td>
</tr>
<tr>
<td>BV89 – percentage of people satisfied with cleanliness of public space</td>
<td>−0.49</td>
</tr>
<tr>
<td>BV90b – percentage of people satisfied with waste recycling</td>
<td>−0.46</td>
</tr>
</tbody>
</table>
the calculated performance for any given authority is not impacted by adjustments to the reported performance of other authorities” (Audit Commission, 2005b, p. 75).

The Commission also listed the seven performance indicators to which adjustments will be made, as listed in table 7 above (p. 76). Yet, given that the Audit Commission endeavours to replace costly and time consuming inspections with performance indicators, and that these PIs are yet to be defined, the question arises whether an adjustment of seven PIs only is sufficient, or whether, in fact, a more structured approach is required as to how to account for deprivation.

The Commission proposes to calculate the adjusted PI value by using the following equation:

\[
\text{Adjusted PI value} = \text{original PI value} - \text{linear function of deprivation}
\]

The function element will use the gradient of the linear regression equation calculated from the 2003/4 data. The Commission also gives a generic example of such a calculation (where the adjustment results in the authority’s performance being increased by 9 percentage points):

Original PI value = 38%
Gradient = -0.25
Deprivation (IMD 2004) = 36
Adjusted PI = Original PI value – (Gradient x IMD 2004)
= 38% – (-0.25 x 36)
= 38% – (-9)
= 47%

4. Avenues for research and preliminary findings

The following questions therefore identify potential gaps in the methodology/objectives of the CPA assessment framework.

1. For the CPA framework in 2005 and beyond, the Audit Commission plans to shift its emphasis away from reliance on inspections towards additional PIs (see section 3.3), such as Best Value Performance Indicators (BVPI). Which impact will this move have on the link between deprivation and CPA scores, given the high correlation between CPA and PIs described in section 2.2?

2. With that greater reliance on performance indicators, are the resulting CPA scores going to be reliable (i.e. results are identical on repeated testing) and valid (the scores truly measure what they purport to measure)?

3. Are there any statistical ‘bounding’ problems, given that during the CPA calculation, performance data are combined, averaged, transgressed and categorised at various stages before the eventual performance score is produced (see section 2.3)?

4. Is there a problem with the Audit Commission using linear regression analysis to correlate data that are not interval, but ordinal (see section 3.4)?

5. Composite measures, such as CPA scores, or for that matter IMD scores, introduce an extra dimension of noise to the data. Are the rankings of local authorities robust to changes in the weighting of the components of their composite scores?

6. Are there predictor variables other than the overall IMD that would better explain CPA results? For example, are there any particular components of the IMD that may themselves be better predictors for particular components of the CPA score (i.e. the scores of the individual service blocks)?

7. Is it possible that the weak correlation between CPA scores and deprivation as attested by the studies stated here is because some factors cancel each other out?

8. Is there an issue with the fact that the IMD scores underlying the Commission’s analysis (section 2.1) and the NPI study (section 2.2) only capture those wards for which deprivation can be attested?

9. More generally, Scottish and/or Welsh approaches to assessing the performance of public service delivery appear to be ‘softer’, allow authorities more flexibility in setting local priorities, and do not impose a similarly uniform assessment framework upon all local entities. Do these approaches have any merit, either in terms of public service performance, efficiency or, in fact, promoting greater ‘localism’?

10. Even more generally, given the compulsory nature of the assessment exercise – with its external inspections; the crude categorisations; the subjection of ‘failing’ councils to punitive interventions; the concessions towards flexibility and freedoms made only if centralist conditions are met; the playing off of
authorities against each other – is it fair to claim that CPA is essentially a ‘centralist’ exercise? Or does CPA promote the very opposite, namely a new localism through the increasing transfer of power (i.e. inter alia financial freedom) to local entities?

Our project is examining as many of these questions as resources permit, except for no. 5, which is being examined by Jacobs et al. (this issue). Statistical findings will be reported as soon as they are available. Some preliminary findings may be of interest.

A new Index of Deprivation (IMD 2004) has been devised by our associates in the Social Deprivation Research Centre, Oxford University and adopted by the Office of the Deputy Prime Minister (ODPM, 2004). Like its predecessor, it reports on the level of deprivation in each of over 30,000 small census units in England, known as Super Output Units. These scores (or ranks) can be aggregated to give scores (or ranks) for each local authority. We are comparing the CPA scores for each council between 2002 and 2004 with these IMD scores, but break each down into finer components: the seven IMD sub-domains of education deprivation; health deprivation; income deprivation; barriers to housing and services; living environment; and crime. We are finding that in some domains (including education, as signalled by previous researchers) high deprivation is associated with a low CPA score. In some others, high deprivation is associated with a high CPA score, giving support to hypothesis 7 above. These results are currently subject to robustness testing, circularity checking and remodelling, inter alia to account for differences.

We have also compared the CPA authorities of English counties and unitaries with their discretionary expenditure. As noted above, each authority is assigned a sum of money, now called Formula Spending Share, which represents the ODPM’s estimate of its relative needs to spend, in order to achieve a standard level of service. Another element of its block grant is supposed to account for the variability of authorities’ tax bases. Local authorities have only one main independent source of revenue, namely Council Tax, which is based on (purported) 1991 house values in eight hands. Authorities in areas of high house prices and/or rapid growth therefore have a more robust Council Tax base than those in poor or declining areas.

If both of these elements of block grant – the ‘needs’ and ‘resources’ elements – measured the real needs and tax bases with perfect accuracy, several consequences would follow. Notably, CPA scores should be unaffected by deprivation, because each council’s level of relative deprivation will have been fully and exactly compensated in its block grant. It might therefore be of interest to correlate authorities’ CPA scores with their discretionary spending, expressed as their actual spending as a proportion of their block grant. This figure ranges from about 0.9 to about 1.3. We find no relationship between these variables, although a weak relationship emerges if one removes a very small number of outlier authorities which have both unusually high CPA scores and unusually low spending. However, we suspect that FSS and the resources element do not produce a perfectly level playing field among authorities.

We have recently finished studying the less prescriptive regimes in Wales and in Scotland, by conducting thirteen semi-structured elite-interviews with senior management in local authorities, auditors, the Scottish Executive, the Welsh Assembly Government, and local government associations. Briefly, we have found that some stakeholders in Wales, among both auditors and auditees, yearn for the more prescriptive English system, as it allows for explicit comparisons to be made between authorities’ performance, a feature that remains unattainable with the flexible Welsh self assessment approach. Yet, we have found no stakeholders in Scotland who do so. The Welsh local government system was closely integrated with the English one until devolution started in 1999. By contrast, the Scottish local government system never has been, its distinctness being asserted in the Treaty and Acts of Union of 1707.

Both Wales and Scotland contain few local authorities (22 and 32 respectively, compared to 150 unitaries and upper-tier authorities in England) and so the relationship between auditor, local government department, and authority can be more intimate than in England: in both nations it is possible to gather all Chief Executives of local authorities in a room and discuss strategy and issues. The production of league tables is rejected for similar reasons: the local authority communities in both nations are small enough for all senior managers to know one another. Putting local authorities in a position of direct competition with one another is therefore not desired. Finally, the small number of authorities brings about assessment frameworks that are marked by a partnership approach between auditors and auditees with a commitment of the former not only to audit but also to support the latter.

The differences between the three nations aside, our work raises two deeper issues: ‘Goodhart’s Law’, and the possibility of contradictory government policies. As is well
known, the ‘law’ originally formulated by Bank of England economist C.A.E. Goodhart in relation to monetary targets, runs, “When a measure becomes a target, it ceases to be a valid measure”. There is much anecdotal evidence in British local government to the effect that performance indicators fall foul of Goodhart’s Law. The decision by the Audit Commission to rely more on them and less on inspections is designed to reduce both subjectivity and transaction costs; but it may worsen the effects of Goodhart’s Law.

To see how government policies may contradict one another, consider education. The CPA regime we studied contains a performance indicator, Percentage of 15 year old pupils in schools maintained by the local education authority achieving five or more GCSEs at grades A*–C or equivalent. The higher an authority’s score on that indicator, the better its chance of a good CPA ranking overall. The Indices of Deprivation 2004 contain a closely related measure for the population of each SOA: Average points score of children at Key Stage 4 (2002). But in this case the relation is inverse. The higher their score, the less deprived is an SOA, and therefore the education authority of which it is part. If the ODPM used the IMD, which it sponsors, directly to determine authorities’ FSS, there would be a direct clash. An omnipotent education officer could either:

- **raise** secondary school performance, get a higher CPA score and more earned autonomy; or
- **lower** secondary school performance, get a higher IMD score and more money for the authority via FSS.

As it happens, the formula for FSS does not use school results directly, so a flat contradiction is narrowly avoided. But the situation repeats itself in many areas of public policy. Should subnational governments be rewarded for poverty, or rewarded for doing well? These two policy objectives can be in severe tension.

**REFERENCES**


