Has Social Care Performance in England Improved?

An Analysis of Performance Ratings across Social Services Organisations

Paul Clarkson, Sue Davies, David Challis (University Of Manchester), Michael Donnelly (Queen’s University, Belfast) and Roger Beech (Keele University)

Please send comments and correspondence to: paul.clarkson@manchester.ac.uk
Has Social Care Performance in England Improved?

An Analysis of Performance Ratings across Social Services Councils

Abstract

Performance measurement in the English public sector has been implemented in a top-down manner. The mechanics of this regime (centralised targets, public reporting of performance data and the use of rewards and penalties) are driven by central government and are intended to hold agencies to account for their performance and regulate their behaviour. This paper examines whether the operation of such a regime has resulted in the improved performance of social care organisations. Comparable data from Northern Ireland Health and Social Services Trusts, where performance measurement processes are less centrally driven and rewards and penalties are less contingent on ratings, are used to investigate the potential impact of different performance regimes. Indicators of admissions to residential care and delayed discharges from hospital are used to signal progress in the meeting of social care objectives in the two countries.

We find that in England social care organisations have broadly improved their performance over time in composite (‘star’) ratings and with generally higher ratings across individual indicators, with noteworthy improvement across key threshold indicators (measures which councils must perform well on to obtain a good summary rating). The comparative analysis highlights the differences across the two countries with an improvement in performance over time in England and a fairly static picture in Northern Ireland. These improvements may reflect a host of factors other than the incentives created by centralised targets and measures. Our data show, for example, that external characteristics of English councils predominantly influence their being reported as ‘poor performers’.

Evidence suggests that the rate of performance improvement reflects the incentives inherent in the English performance system. An incentive structure that promotes the
achievement of specific targets may have influenced the behaviour of local managers responsible for performance in this sector. However, further information from local organisations is required to more fully explicate all the relevant factors that may have a bearing on reported performance. The evidence reported here does, however, point to the beneficial effects (at least for the regulator) of centrally set measures. It is uncertain whether more recent moves away from this target regime in England, towards the development of local indicators and ‘individual incentive’ regulation, will likewise offer the appropriate inducements for performance improvement.

Acknowledgements

We would like to thank members of the North West Performance Leads group in social services in England, and in particular David Burnham, for their helpful advice and assistance with the wider project from which the data here relates. We also thank Gary Bennett of the Northern Ireland Department of Health, Social Services and Public Safety for making available the data on delayed discharges. PSSRU also receives funding from the English Department of Health and the views expressed here are those of the authors and not necessarily those of the Department of Health, the ESRC or other fundors. This is a joint ESRC Public Services Programme and PSSRU paper and therefore also appears in the PSSRU Discussion Paper series as paper no M188.
1. Introduction

Performance measurement has been implemented in a range of public services since the 1980s and a number of aims have been ascribed to it (Propper and Wilson, 2003). Among these has been a wish to uphold a greater degree of accountability in the way local agencies, such as hospitals and local authority departments, deliver their services; the need for the public and those who use services to receive information by which they can exercise choice and to have the right of redress if anything goes wrong; and the requirement for managers to manage their services with the assistance of a range of performance data locally. These systems of performance measurement have been vigorously applied to public services in the UK, the United States and other countries (Kravchuck and Schack, 1996; Propper and Wilson, 2003; Bloomfield, 2006).

In the UK public sector, and more so in England as opposed to other countries of the UK, performance measurement has been implemented in a particularly top-down manner (Sanderson, 2001). Performance regimes, applied to education, health care, local government services and others, have been clearly associated with the setting of standards and targets, and their linkage to the public reporting of performance where rewards and punishments are given to endorse or admonish ‘good’ or ‘poor’ performance respectively (Bevan and Hood, 2006a). Here, the measurement of performance is tied to a set of incentives with resulting processes such as ‘earned autonomy’ for those performing well and ‘naming and shaming’ for those performing poorly, with resulting powers to impose sanctions (Huber, 1999; Guardian, 2001). The mechanics of these regimes, in England at least, are driven by central government and are intended to regulate and monitor the behaviour of local agencies such as NHS Trusts and local authorities.

A question for policy makers is whether the operation of such regimes has resulted in performance improvement. Several accounts provide a context for exploring this question. First, there are complexities involved in seeking to improve performance and even defining what such improvement would look like. The initial design of measures will therefore affect how performance is judged. Behn (2003), for instance, draws attention to the multiple purposes of performance measures and how, for
improvements to take place, the measures chosen must be appropriate for particular purposes. If the purpose is to control, then measures of inputs are needed (such as, for social care, the numbers of home care staff) as these can specifically be regulated in terms of whether policy requirements have been adhered to and often lie within the responsibility of managers to commission or provide. If, on the other hand, organisational learning is the key purpose behind measuring performance (Garvin, 1993), then a wider set of measures is needed. These measures would include more detailed information on the processes that occur within organisations and the compliance of these with ‘best’ professional or organisational practice (Challis et al., 2006a). In many regimes, however, the purposes behind measuring performance are rarely stated explicitly and multiple purposes are often hidden behind a single set of measures. In centrally-driven regimes, these measures are, to an extent, imposed on local agencies, which can create incentives to respond to the measures themselves not the intention behind them (Smith, 1995). Second, following from this, the incentive structures to promote improved performance must be carefully considered. This is the issue of how to motivate agencies to improve through the publication of information. Dixit (2002) has pointed to the divergent means by which control by the ‘principal’ (government or regulator) towards ‘agents’ (regulated agencies such as local authorities) may be affected through incentive schemes where asymmetries of information exist. His analysis suggests that, where there are multiple sources of information for the principal to observe (such as performance measures reported to government), and multiple actions involving multiple principals (for example, government in its relationship with local management and managers’ relationship with front-line staff), then a performance-control system must be designed carefully to minimise gaming behaviour. In this sense, incentives are required that offer payoffs to both principal and agent for improving performance. Improvement is therefore not a simple matter; it is dependent on a host of factors associated with the design of performance systems and the responses of local agencies to them.

These accounts offer a framework for considering the complexities involved but there is little evidence that the operation of these performance regimes per se has resulted in better performance (Propper and Wilson, 2003). Advocates of these regimes argue that a focus on national targets has led to improved performance (Audit Commission, 2003). But, performance improvement has been viewed as contingent on a variety of
circumstances within organisations in addition to the effects of incentives arising from control and reporting systems (Scott et al., 2003; Boyne and Enticott, 2004; Mannion et al., 2005; Jas and Skelcher, 2005). Moreover, there are measurement issues to consider when examining performance information, such as the error that may be part of reporting data and regression to the mean and other statistical artefacts of measurement over time (Propper and Wilson, 2003; Greve, 1999). These issues tend to confound attempts at attributing improvements in performance to policy changes alone. The above accounts do, however, point to the need to consider the specific characteristics of performance schemes and the different incentives offered for agencies to improve their performance. Several questions therefore arise: are there particular elements within these systems that act as a spur to improved performance? Are there lessons to learn from comparing performance regimes in operation across different settings, such as health, education, or social care, or between different countries where these regimes operate slightly differently?

This paper explores reported performance in English social care organisations as a means of investigating these questions. Social care, administered through local authorities, is a locally delivered service operating to centrally determined policy goals. It thus provides a setting in which to examine the effectiveness of centrally-driven policies for performance improvement. These policies, in England, have become more pronounced over the last decade, with central control and the imposition of national standards far more in existence. The relative balance between centrally set and independently designed standards for local authorities has been an issue for much of the post-war period (Banwell, 1959). There is therefore scope for exploring the effects of performance regimes in which this balance is differently configured. The effects of ‘naming and shaming’ social services councils may, for example, not be the same across the UK as that reported for the centrally-driven NHS (Bevan and Hood, 2006b). Different countries of the UK and, in fact, English practice viewed historically, also exhibit differences in approach (Haubrich and McLean, 2006). The hypothesised causal link between the use of comparative measures of performance and actual performance improvement (often only stated implicitly) may also be confounded in social care by the influence of local circumstances and even resistance to central government monitoring. The goals of a locally delivered service may be at odds with those of central government or, in social care, may be more difficult to
define clearly, thus raising issues for measuring the practices of those subject to performance monitoring, local managers. In the following analysis we consider the performance of English social care on a number of indices including composite (‘star’) ratings and key indicators. We also examine the characteristics of ‘poor performers’ to ascertain whether local exogenous circumstances, such as deprivation or a potentially high demand for services, may partly explain their limited performance, as a way of teasing out these issues.

In this analysis we also compare English data with that from Northern Ireland, where performance measurement for social care organisations (in that country integrated Health and Social Care Trusts) operates slightly differently. In Northern Ireland, there has been no system of explicit government targets, nor have authorities been ‘named and shamed’ in terms of their comparative performance. These countries are analysed on two comparable indicators, which express broad objectives of social care: admissions to residential/nursing care and delayed discharges from hospital. A comparison of performance ratings between these countries of the UK is likely to offer insights as to the way measures are used and the potential impact of different performance regimes.

To provide a context for the analysis, we first describe, in the following section, the use of performance measures in social care and draw on this to establish five hypotheses regarding performance, principally in England, particularly for older people who form the bulk of service recipients. We then briefly outline our methodology and data analysis. Finally, we report our findings and conclude by considering a number of policy implications for the future practice of performance measurement in this sector.

2. **Performance Measures in Social Care**

Whilst there have been sets of performance indicators in English social care since the mid to late 1980s (see Miller, 1986; Warburton, 1988), it is only since the modernisation reforms of 1998 onwards that performance ratings on key indicators have been tied explicitly to the regulation and control of local authority social services councils (Cm 4169, 1998). From that time, sets of nationally published indicators
have been used to establish standards for the conduct of English social care. Summary measures (‘star’ ratings), rated across councils, have been published from May 2002 (Social Services Inspectorate, 2002). The stated aim behind these ratings was to examine variations across councils and to signal where those councils performing poorly needed to improve (see Commission for Social Care Inspection, 2006a). The rationale behind this, rarely stated explicitly, was one of ‘yardstick competition’, where comparative performance ratings were used as a discipline for local councils to ‘mimic’ the behaviour of their peers (Shleifer, 1985; Ladd, 1992; Revelli, 2006); in the words of the 1998 White Paper Modernising Social Services, authorities were to use these comparative ratings ‘to drive up their standards to match those of the best’ (Cm 4169, para 7.3). Within this rationale, much of the variation in service delivery was considered to be unacceptable and other potential sources of variation, such as the quality of the data, how measures were calculated, and the resources available to local authorities were largely considered irrelevant to an assessment of councils’ performance (see Milburn, 2001).

The incentives generated by this performance system have been associated with a number of behaviours amongst local social services managers. Whilst these are not well documented, several local accounts see the effects of these ratings as galvanising local managers to monitor their local processes more explicitly in order to safeguard their reputations, whilst the effects for staff on the ground, such as social workers, have been ones of distrust and loss of morale (Burnham, 2006; Caldwell, 2003). Other accounts in the social care professional press have been more distrustful. National ratings have been characterised as a simplistic attempt to capture the range of social care provision. The use of comparative ratings to reward or punish good or poor performance respectively has also, it is argued, largely neglected the often difficult circumstances facing ‘poor performers’, which made it difficult for them to match the performances of the best (Revans, 2002; Cutler, 2002). The question arises, therefore, as to whether incentives to improve performance through national targets are stronger than those operating for local decision-makers (such as social workers) who largely generate the data used for performance review and who view national targets as too blunt an instrument for capturing the variety of constraints on their practice.
In England, the measures chosen to signal improved performance have been centrally set, and more recently, have been disseminated at arms length from government through the main inspection body for social care, the Commission for Social Care Inspection (CSCI). These measures have included both composite ‘star’ ratings for councils and also a set of ‘Performance Assessment Framework’ (PAF) indicators for adults, which are used to assess the variation across authorities on a range of measures aligned with broad objectives, such as those to care for people in community rather than institutional settings or to promote the timely delivery of services (Commission for Social Care Inspection, 2006b). Different standards have been set in order to produce a range of ratings from these measures. A sub-set of measures, termed ‘Key Threshold’ indicators have also been contained within the PAF indicators. Key Threshold indicators are those held to be important judgments of performance by central government (Commission for Social Care Inspection, 2006a). Failure to meet centrally defined levels of performance against these indicators negatively affects the overall composite performance (‘star’) rating, across authorities, published each year. Over and above these regulatory indicators, a wider set of measures (‘key indicators’) is also available which authorities themselves, inspection bodies and researchers may use. However, this performance system for authorities is set to be differently configured. A new, reduced National Indicator Set has been proposed as one part of a ‘lighter touch’ from central government and as a mechanism by which local authorities can measure their performance (DCLG, 2007a). The previous national systems – the PAF set of nationally prescribed measures for all social services councils and the composite Star Ratings – are to be abolished and this is set to effect the 2009 judgements of local councils (DCLG, 2007b; CSCI, 2008). However, some statutory targets are likely to remain, albeit in negotiation with localities through Local Area Agreements (LAAs). It is not yet clear exactly what information will be available specifically for social care or what, if any, targets will be set.

In Northern Ireland, Key Indicators of Personal Social Services (Northern Ireland Statistics and Research Agency, 2007) are used to compare the performance of Trusts and facilitate the monitoring and review of individual local services as well as being

---

1 These measures are available on the Dr Foster Intelligence site at http://www.drfoster.co.uk/localGovernment/kigs.asp.
used to generate analytical reviews of community care by, for example, the Social Services Inspectorate. Each Health and Social Services Trust has its own system for monitoring performance activity and the Key Indicators are used, in the main, to offer a descriptive account of provision. In Northern Ireland there has been no direct equivalent of Star Ratings for social care, as in England, and, overall, performance measurement processes appear to vary more and to be less centrally driven. The performance system has comprised a series of regular meetings, reports and reviews between the organisational units at each tier of Northern Ireland’s health and personal social services, based on annual plans and performance goals set out by the Department of Health, Social Services and Public Safety (DHSSPS) in Priorities for Action documents. This system has been criticised in terms of its lack of information about the achievement of targets over time and about the consequences following from meeting or missing targets and goals (Appleby, 2005). As a consequence, Northern Ireland’s integrated health and social services system has recently undergone a process of major reorganisation as part of the review of the management and delivery of public services (DHSSPS, 2008). This reorganisation is to implement new structures from 1 April 2009 and is to also offer a sharper focus on performance management, in particular the meeting of targets. The DHSSPS will become smaller although one of its roles will be to set strategic targets and monitor outcomes for the reorganised service.

It is instructive to analyse and compare the performance of these units in both countries so as to explore what may drive performance improvement and whether the incentives, or lack of them, in operation may be associated with particular outcomes. This would seem to be especially the case since the systems in the two countries of the UK seem set to change, albeit in opposite directions; there is to be a contraction in the range of regulatory measures in England along with a ‘lighter touch’ from central government and fewer central targets, whereas in Northern Ireland a greater degree of systematic targets and tighter performance monitoring have been called for. The basis of a regulatory use of performance measures, as has occurred in England, is one of ‘regulation for quality’ (Laffont and Tirole, 1993) and there may be a host of beneficial, and otherwise, outcomes that may result. Given this rationale for national performance ratings, and the difficulties raised by local practitioners, what have been the likely impacts on the performance of social care? The following hypotheses are
stated as reflecting expected changes from the implementation of performance measurement in social services:

- summary performance across English social services councils will have improved over time with the proportion of councils receiving higher ratings increasing.
- the majority of English PAF indicators will have shown improvements in performance over time.
- where performance has been stressed in particular areas, through designation of a sub-set of indicators as ‘key threshold’ indicators, performance will have shown improvement in these areas as against other areas in which there will have been more variability.
- performance in Northern Ireland, where the discipline effect of ratings has not existed, will have shown little or no improvement compared with that across England.
- there will have been significant associations between authority characteristics and designation of authorities as ‘poor performers’ in England.

As a way of drawing out some of the above issues, the following analysis uses data (composite measures and PAF indicators) to explore, in a retrospective fashion, the extent of improvement in reported performance across social services councils in England and compares two indicators (residential admissions and delayed discharges), where these present comparable data, from Northern Ireland. This analysis was conducted to form a picture of the situation at a particular point in the development of performance measurement in social care in the two countries (late 2007/early 2008).

3. Data and Analysis

The analysis conducted here formed part of a wider project on performance measurement in local social care organisations funded by the Economic and Social Research Council’s Public Services Programme (Clarkson et al., 2007). Data for all English social services councils were interrogated from the year 2002, when summary measures (‘star’ ratings) became available (Commission for Social Care Inspection, 2006a,b, 2007a). Where comparable data were available, these were compared with
those from the performance system in Northern Ireland, which has not adopted a star ratings system and where performance data are used instead to provide “detailed comparative information that affords Trusts and Boards the opportunity to compare service provision and expenditure with that in other areas…to help to raise questions about the complexity of social service provision” (Northern Ireland Statistics and research Agency, 2006a, p.19).

The units of analysis here were the councils with social services responsibilities (CSSRs) in England, of which there were 150, comprising diverse populations and circumstances from large metropolitan councils, such as Manchester, to counties, such as Lancashire, the London Boroughs and the newer unitary authorities, such as Warrington. In Northern Ireland the units with similar responsibilities were the integrated Health and Social Services Trusts, of which there were 11. From April 2007 the review of health and social services provision in Northern Ireland reduced the number of these units down to five (Appleby, 2005). These units were: Belfast, South Eastern, Southern, Northern, and Western Health and Social Care Trusts.

We first compared the national star ratings in England in terms of the number of councils receiving 3 stars (‘excellent’), 2 stars (‘good’), 1 star (‘adequate’) and zero stars (‘inadequate’) for each year from 2002-2007. We then subjected each of the PAF indicators for social services specifically for older people in England (Commission for Social Care Inspection, 2006b), of which there were 20, to the same comparisons over the same time period. These were indicators of Adults Services, being predominantly for older people over 65 years, the greatest users of social services. These indicators described authorities’ positions on a range of performance dimensions, including: the timeliness by which assessments and services were provided; service delivery, such as numbers of older people entering residential forms of care or receiving care at home; quality indicators (such as the number allocated single rooms in residential care); costs; and compliance with central government objectives, such as the numbers receiving direct payments (cash payments made to

---

2 Key Indicators of Personal Social Services for Northern Ireland available at: http://www.dhsspsni.gov.uk/statistics_and_research-cib-key_indicators.
individuals who have been assessed as needing services, in lieu of social service provision) (Department of Health, 2003).

Summary measures are not available for social care in Northern Ireland with most indicators reflecting outputs, i.e. numbers of people receiving specific types of services such as home care or residential care, rather than outcomes (such as satisfaction ratings) or processes (such as waiting times for the delivery of assessments or care packages), which are available in the English package. Indicators in Northern Ireland are also not subject to performance ratings against a nationally imposed standard as they are in England. Targets have been used, but these have tended to be either erratic or with no well defined milestones for achievement and few sanctions for non-achievement (Appleby, 2005). Therefore, when comparing Northern Ireland with England, raw scores were used (i.e. without them being subjected to performance ‘rating’) and, where appropriate, these were standardised according to the same denominator value, for example per population aged 65 and over. For some of these indicators, data were available from before 2002, when the summary ratings of performance in England were introduced, permitting a comparison of the different incentive systems in operation in the two countries.

Northern Ireland was compared with England on indicators where these presented comparable information. These were almost exclusively related to outputs, such as the form of provision across authorities/Trusts. Home care, delivering personal and domestic care in people’s homes, was defined differently in the two countries with English figures concentrating on ‘intensive’ home care (defined as more than 10 contact hours or 6 or more visits during a sample week). Indicators of numbers receiving home care in general were therefore likely to be underestimates in England and so were not directly comparable. Residential homes tend to be considered as a separate form of provision from nursing homes but in both countries information on the numbers admitted to these are combined. This indicator (numbers admitted to residential/nursing homes per 10,000 older people over 65 years) therefore provided a comparable measure of whether social care objectives were being achieved as the proportion of people using this form of provision might be expected to stabilise or decline over time as more people are being cared for at home. However, the ‘balance of care’ literature shows that the numbers admitted may depend on a host of other
factors, not least the supply of residential places (Challis and Hughes, 2002; Clarkson et al., 2005). To account for this, an adjusted indicator was also compared between the two countries allowing for the number of places available for the local authority/Trust to purchase or provide. Numbers of residential places in Northern Ireland are contained in the Key Indicators package (Northern Ireland Statistics and Research Agency, 2007) and in the Community Statistics publication (Northern Ireland Statistics and Research Agency, 2006). For England, comparable figures are no longer available as key indicators and so these data were drawn from the Department of Health’s Community Care Statistics (Department of Health, 2001) and the Commission for Social Care Inspection’s *State of Social Care in England* report (CSCI, 2006c).

Waiting times for inpatients and outpatients has been identified as a particular problem in Northern Ireland (Appleby, 2005) and here the delay in discharges from hospital presented a comparable measure in the two countries. In England, delays in hospital discharge (where the patient is awaiting arrangements for care at home or in residential/nursing care through social services) have been monitored through the ‘Delayed Transfers of Care’ indicator, which measured the number of delayed transfers from acute hospitals per 100,000 population aged 65 and over (Commission for Social Care Inspection, 2006b). This indicator was designed to signal progress in the extent of delayed discharges to monitor the success of the Community Care (Delayed Discharges etc.) Act 2003, where social services councils were to pay NHS bodies a charge for delays where social care arrangements on discharge had not yet been made (Department of Health, 2003). In Northern Ireland, the provisions of the Act do not apply but data have been collected on discharge delays since 2000. These data, sourced from the Delayed Discharge (DD1) return to DHSSPS, are collected by monthly census for each hospital and returned centrally. For this analysis, each set of data was calculated, on average, per 100,000 older people to permit comparison using population estimates for each country.

Finally, analysis was undertaken to examine the external characteristics of English local authorities and their relationship with performance ratings. This analysis was to test the proposition that ‘poor performers’ are largely so due to difficult circumstances rather than ‘bad choices’ by managers whose performance is being monitored
(Andrews et al., 2005). Four variables were considered that reflect the quantity of service need and prosperity of residents following the work of Andrews et al., (2005): percentage of population aged over 65, percentage of pensioners living alone, percentage of households with at least one person of pensionable aged in rented accommodation and the amount of formula spending share (FSS) in proving personal social services that is spent on the over 65s\(^3\) In order to examine the characteristics of ‘poor performing’ authorities, analysis of variance tests were undertaken examining the association between these variables and star ratings. When a simple one-way ANOVA test was not appropriate (due to assumptions of homogeneity or normality not being met) the Kruskal-Wallis test was applied.

4. Findings

Figure 1 shows the composite ‘star’ ratings for all English social services councils from 2002 (when the ratings were introduced) to 2007. There has been a tendency for the number of councils receiving higher ratings (2 and 3 stars) to increase over time and for the numbers receiving lower ratings (zero and 1 stars) to decrease. Around 81% of councils received the highest ratings of 2 or 3 stars in 2007 compared with just 33% in 2002.

Figure 1: Composite ‘star’ ratings for all English social services councils from 2002 to 2007

\[\text{Source: Commission for Social Care Inspection (2006a).}\]

\(^3\) The FSS is calculated in order to reflect the spending needs of an authority. The formula takes into account information regarding the demographic, social and physical features of each authority to produce an assessment of financial need and replaces the previous Standard Spending Assessment.
Table 1 shows the performance ratings for English social services councils across all 20 ‘key indicators’ used to form national judgements of social care activity. There were improvements across the board according to these ratings for most of the indicators. However, where indicators were classified as ‘key thresholds’ they universally showed improvements whereas the picture exhibited by other indicators was more mixed. In particular the indicator ‘cost of intensive social care’, a measure of the efficiency with which social care councils deliver services showed little improvement over time with, in fact, a deterioration since 2006 in the proportion of councils receiving the highest ratings.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>C51 Direct payments *</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>35</td>
<td>61</td>
</tr>
<tr>
<td>D54 Timely delivery of equipment *</td>
<td>-</td>
<td>-</td>
<td>79</td>
<td>81</td>
<td>97</td>
<td>98</td>
</tr>
<tr>
<td>D55 Timely delivery of assessments *</td>
<td>-</td>
<td>-</td>
<td>33</td>
<td>37</td>
<td>53</td>
<td>48</td>
</tr>
<tr>
<td>D56 Timely delivery of services *</td>
<td>-</td>
<td>-</td>
<td>80</td>
<td>81</td>
<td>87</td>
<td>81</td>
</tr>
<tr>
<td>C62 Carers receiving a service</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>46</td>
<td>71</td>
</tr>
<tr>
<td>C72 Admissions to residential care</td>
<td>-</td>
<td>-</td>
<td>71</td>
<td>79</td>
<td>86</td>
<td>93</td>
</tr>
<tr>
<td>B11 Balance of home/residential care</td>
<td>-</td>
<td>-</td>
<td>71</td>
<td>79</td>
<td>86</td>
<td>93</td>
</tr>
<tr>
<td>B12 Cost intensive social care</td>
<td>-</td>
<td>43</td>
<td>49</td>
<td>41</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>B17 Unit cost home care</td>
<td>35</td>
<td>52</td>
<td>47</td>
<td>45</td>
<td>37</td>
<td>45</td>
</tr>
<tr>
<td>C26 Admission to residential/nursing care **</td>
<td>65</td>
<td>75</td>
<td>87</td>
<td>92</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C28 Households intensive home care**</td>
<td>38</td>
<td>43</td>
<td>45</td>
<td>50</td>
<td>57</td>
<td>59</td>
</tr>
<tr>
<td>C32 Older people helped at home</td>
<td>45</td>
<td>40</td>
<td>39</td>
<td>38</td>
<td>43</td>
<td>41</td>
</tr>
<tr>
<td>D37 Those allocated single rooms</td>
<td>69</td>
<td>77</td>
<td>77</td>
<td>82</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>D39 Those receiving statement of needs</td>
<td>18</td>
<td>20</td>
<td>29</td>
<td>39</td>
<td>59</td>
<td>75</td>
</tr>
<tr>
<td>D40 Proportion receiving a review*</td>
<td>-</td>
<td>-</td>
<td>48</td>
<td>69</td>
<td>79</td>
<td>92</td>
</tr>
<tr>
<td>D41 Delayed transfers of care</td>
<td>-</td>
<td>70</td>
<td>82</td>
<td>94</td>
<td>98</td>
<td>79</td>
</tr>
<tr>
<td>D52 Satisfaction with services</td>
<td>-</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>51</td>
<td>-</td>
</tr>
<tr>
<td>E47 Ethnic minorities assessed1</td>
<td>23</td>
<td>36</td>
<td>42</td>
<td>56</td>
<td>66</td>
<td>72</td>
</tr>
<tr>
<td>E48 Ethnic minorities receiving services1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>57</td>
<td>69</td>
<td>76</td>
</tr>
</tbody>
</table>


* Key Threshold indicators
** Deleted as Key Threshold indicator in 2006
- Indicator not used.

In terms of a comparison of outputs between England and Northern Ireland, Figure 2 shows the numbers of older people admitted to residential/nursing homes from 2000/01 to 2004/05. The numbers admitted in Northern Ireland, on average, stayed fairly steady at around 103 older people per 10,000. Those in England, however, showed a fall after 2001-02 from around 111 per 10,000 to 90 per 10,000 (a reduction
of over 20%). In terms of the adjusted indicator, which takes account of the number of residential or nursing home places available, England again showed a consistently lower figure than Northern Ireland, with around 4% (per 10,000) of available places being occupied. The fall in rates since 2001-02 in England was not as dramatic as the raw figures on the number of admissions suggested. Rather, there was a gradual decline in the numbers of admissions as a proportion of the places available (Figure 3).

**Figure 2:** Comparison of numbers of older people admitted to residential/nursing homes in England and Northern Ireland from 2000/01 to 2004/05

![Figure 2](image1)

**Sources:** Commission for Social Care Inspection (2006b); Northern Ireland Statistics and Research Agency (2006a); Numbers of places in both residential and nursing homes are included to illustrate total residential provision in Northern Ireland. As a consequence residential beds in nursing homes have been excluded to avoid double counting.

**Figure 3:** Comparison of numbers of older people admitted to residential/nursing homes in England and Northern Ireland as a proportion of number of places available

![Figure 3](image2)

**Sources:** Commission for Social Care Inspection (2006b); Northern Ireland Statistics and Research Agency (2006a); Department of Health (2001); Commission for Social Care Inspection (2006c).
For delayed discharges, comparable figures between England and Northern Ireland showed consistently lower numbers of delayed discharges in England. On average, the numbers per 100,000 whose discharges were delayed fell in England from 2002-03, whilst those in Northern Ireland remained fairly static until 2004-05 when they began to fall (Figure 4). However, this comparison of average figures for the two countries masks considerable local variation. In England, councils ranged from between 0 to 93 (per 100,000) delayed discharges for 2005-06. In Northern Ireland, the review by Appleby (2005) described the numbers of patients with delayed discharge as fairly static, remaining at around 350-400 in any one month between 2000 and 2004; this is more than twice as high as in England. However, this figure again may mask local variation with certain Trusts showing relative improvement after discharge delays were accorded priority (South Eastern Health and Social Care Trust, 2008). Moreover, differences in delayed discharges may simply reflect differences in hospital admission rates between the two countries (with changes in the proportions of hospital admissions that suffer delays therefore being a more reliable indicator). However, Appleby (2005) concludes that Northern Ireland has had a similar level of aggregate hospital activity to England, which would suggest that these differences in delayed discharges are a valid reflection of the processes in the two countries.

**Figure 4:** Comparison of number of delayed discharges in England and Northern Ireland 2002/03 to 2005/06 (average figures)

Sources: Commission for Social Care Inspection (2006b); Northern Ireland Delayed Discharge (DD1) return.
In terms of the characteristics of local authorities designating poor performance, in England the three variables; percentage of pensioners living alone, percentage of households with at least one person of pensionable age in rented accommodation and the amount of FSS spent on over 65s, were found to be significantly related to star rating. As shown in Table 2, the mean rank for each of these three variables increased with star rating; the higher the measure of need, the lower the star rating. When interpreting the findings displayed in this table it is important to remember that the Kruskal-Wallis test uses a mean rank in which cases are ranked from 1-150, with 1 being the authority with the most need and 150 the authority with the least. A low mean rank therefore represents a grouping of higher need; for example, authorities with the highest percentages of pensioners living alone will receive a higher ranking, which in turn translates to a lower mean.

Table 2: English performance ratings with mean (ANOVA test) or mean rank (Kruskal-Wallis test) by external variable

<table>
<thead>
<tr>
<th>Performance 'star' rating</th>
<th>Aged over 65 (mean)</th>
<th>Living alone (mean rank)</th>
<th>Rented accommodation (mean rank)</th>
<th>FSS (mean rank)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Star</td>
<td>16.20</td>
<td>59.66</td>
<td>62.02</td>
<td>61.66</td>
</tr>
<tr>
<td>2 Stars</td>
<td>15.51</td>
<td>74.17</td>
<td>73.01</td>
<td>72.43</td>
</tr>
<tr>
<td>3 Stars</td>
<td>14.46</td>
<td>86.79</td>
<td>87.01</td>
<td>88.31</td>
</tr>
</tbody>
</table>

* Significant at 0.05 level

Sources: Key Indicators Graphical System (KIGS).

5. Conclusions

Measuring performance in social care, in common with other settings, is a complex endeavour, not least because of the multiplicity of objectives, the range of measures available and the multiple stakeholders that performance information must satisfy. In this context, performance measurement in England has been implemented in a particularly top-down manner with control by central government as a principal aim. The nature of this control contains potential for tension between government and the local authorities administering care services (Banwell, 1959). The system which has developed in England has arisen out of this tension and, in line with developments in other sectors, has sought control through the presentation of comparative information.
The intention has been one of regulating local social services councils through the discipline effect of performance reporting (‘yardstick competition’). In this paper we have examined the trends in reported performance for English social care in order to raise issues concerning whether this aim has been achieved. These trends, at least partly, reflect the incentives inherent in this system and to this end we find support for all our hypotheses.

Our findings show a marked tendency for summary measures of performance, ‘star’ ratings, to have improved over time – what has been termed ‘star creep’ (Burnham, 2006). However, as others have suggested, such apparent improvement may reflect the fact that:

- performance has actually improved and reported performance accurately reflects ‘true’ performance;
- reported performance has been accurate but that this has masked poor performance on the ground in areas where performance has not been measured;
- reported performance has improved but this only reflects the responses of those reporting the data and their unintended gaming behaviour (‘hitting the target and missing the point’);
- ‘true’ performance has not improved but that this has been concealed by unreliability of the data or outright fabrication which has tended to show reported performance in a good light (Bevan and Hood, 2006b).

Disentangling such sources of improvement from the data reported here is difficult. Whilst summary measures have shown steady improvement for councils, the picture for specific indicators is more mixed. Those measures which simply indicate the achievement of central government policy priorities at the local level (for example the numbers receiving direct payments) have shown substantial improvement. Others, such as ‘the cost of intensive social care’, an efficiency measure (Clarkson and Challis, 2006); have shown little or no reported improvement over time. Similarly, performance on the key threshold indicators (measures on which councils must have performed well to obtain a good summary rating) has, on the whole, improved whilst
that on some other measures has not, giving some credence to the view that where performance is not measured (or in this case, an aspect of it is not stressed to as great a degree) then it will not be given the same priority by local officers reporting the data (see Ridgway, 1956 for an early example of this issue). Such behaviour is not surprising given the incentive structures in operation and it may therefore be the case that poor performance on the ground has suffered in areas where performance has not been measured. Some local accounts in social care also point to the fact that, whilst summary performance has improved, performance from the point of view of front line staff in some areas has not (Community Care, 2003). Evidence that gaming behaviour by senior managers and information officers in social services councils has fed directly into reported performance improvements is slight compared with that in other sectors such as the NHS (Goddard et al., 2000) and education (Wilson et al., 2006). However, accounts by Burnham (2006) and Miller (2004) point to behaviour by senior managers, such as informal quota systems for social workers assessing older people directed at reducing overall the numbers in residential care, which may have ‘hit the target and missed the point’. Concealing performance decline by reporting data that are of poor quality has previously been suggested to be a problem for social care councils (Burnham, 2006) but the extent of this across England, and particularly the existence of outright fabrication are unknown.

To fully examine the relative weight that may be ascribed to each of these potential sources of apparent improvement requires knowledge of: local performance management processes; the behaviours of local social services managers in response to the monitoring regime that has emerged; and the specific managerial strategies employed to improve the quality of social care delivery. These behavioural factors, which may be largely responsible for improving performance, operate between the work of front-line staff (whose activity largely forms the data on which performance management is based) and the measurement and reporting of performance (de Waal, 2003; Yang and Hsieh, 2007); these factors are managerial rather than political (Behen, 2003; Melkers and Willoughby, 2005), local rather than national. To investigate these factors and their relative influence on ratings requires a study of local councils as, currently in social care, we simply do not know how managers across England make use of data locally and the extent to which their behaviours
To develop this further, our analysis focuses only on reported performance and trends. The analysis cannot isolate the particular factors that may have accounted for these improvements and we cannot infer that these improvements are necessarily reflective of practice on the ground or whether they can be directly attributed to the incentives created. For this, multivariate analyses are necessary and the wider project we are conducting involves analysis of the predictors of performance ratings and the particular arrangements that are associated with successful monitoring. However, the trend towards improvements in reported performance in England concurs with that in other sectors. In the NHS, for example, there have been improvements, chiefly in waiting times, reported in England compared with other countries of the UK (Alvarez-Rosete et al., 2005; Bevan and Hood, 2006b). In wider local authority services, overall improvements have been reported. Increases in Comprehensive Performance Assessment (CPA) star ratings have been reported by Grace and Martin (2008); the number of authorities receiving a three star rating or above increased by 11% from 2005 to 2006 and in 2006 not one authority achieved the lowest possible performance assessment rating. Their analysis of indicators by service area provides some evidence to suggest improvements have been seen in all service areas from 2001 to 2006, albeit judged by performance over a small number of indicators. The rate of improvement does, however, vary across service areas. These accounts draw on data arising from ‘natural experiments’ in post-devolution Britain, which all appear to point to the positive consequences (at least from the regulator’s viewpoint) of adopting centrally-set targets in England compared with other UK countries.

A further feature to understanding the design of performance systems in this and other sectors lies in the comparative exploration of performance measures in Northern Ireland compared with England. The absence of summary ratings and explicit government targets in Northern Ireland makes for a potentially fruitful comparison in

---

4 A survey of local councils, capturing their management arrangements, use of local performance data and procedures for designing and collecting relevant performance information has been administered in England and Northern Ireland as part of the wider project. The data from these surveys are, at the time of writing, being analysed.
terms of what may drive performance improvement. Our findings suggest that, in terms of the central objective to care for older people outside of institutional settings, the picture in Northern Ireland is more muted than that in England. Delayed discharges from hospital of patients needing care by social services has also been signaled as a problem in Northern Ireland and this is despite closer links between health and social services (Challis et al., 2006b). There may be lessons from other sectors here. For example, performance improvements in the NHS in England are reported to have had a discipline effect on the NHS in other UK countries (principally through the waiting list targets) (Bevan and Hood, 2006b). Such an effect appears to apply in social care with the discipline of national targets now being advocated in Northern Ireland to promote improvement (Appleby, 2005). Therefore, although in England the role of national targets and prescribed measures has been questioned, the current evidence suggests that there may still be a role for government targets as a spur to improved performance (Hood, 2007).

Other factors are at play, however, and the data reported here also suggest influences on performance ratings not considered in the original articulation of the regime. Our findings show, for example, associations between certain external characteristics of English authorities and a tendency to be rated as poor performers; findings that are congruent with other local authority research (Andrews et al., 2005) and cast doubt on the capacity of national regulatory regimes to fully explain improvement solely by reference to the choices and strategies of local managers. Therefore, the general theory of improvement lying behind performance measurement regimes appears constrained by exogenous factors that have an impact on authorities’ performance and capacity to change.

Of interest for future research is the fact that this analysis was conducted at a particular time in the development of policies for performance measurement in this sector. The analysis here presents a view of performance improvement for English councils after a sustained period of centrally-prescribed targets and measures. The picture of improvement indicated herein, although with the important caveats above, reflects the incentives contained within this system. At a time when the emphasis on national targets and control has waned somewhat, it is not yet clear just how the emerging moves towards localism and recognition of the need to monitor local
capacity and improvement (DCLG, 2008), will work through in practice. In particular, it remains uncertain just how such a system of ‘individual incentive regulation’ (Chong, 2004) will offer the appropriate inducements to improve or even whether the necessary local information can be provided (Warburton, 1999; Challis et al., 2006a). Longitudinal research could examine these issues by tracking indices of performance improvement through time and by observing the effects of such policy changes. The literature testifies to the importance of viewing performance measurement within such an evolutionary perspective, with reported performance being seen as due, in part, to factors such as gaming over time in addition to the desired changes in behaviour of those being regulated (Courty and Marschke, 2003).

It appears that these proposed changes in England (which are still under consultation) are in line with changes in other sectors. In wider local authority services, for example, national prescriptive systems have been replaced with systems of self assessment, where local authorities present their own pictures of performance, through data collected at the local level (Haubrich and McLean, 2006). Similarly in social care, self assessments have been undertaken since 2007 and are currently playing a wider role in the generation of performance judgements (CSCI, 2008).

There also seems to have been a recent convergence in the policies adopted in different countries of the UK, with moves towards fewer national targets in England but with a call for a greater degree of formal regulation against standards in the other countries of the UK (Blackman et al., 2006).

The pattern of changes, although conducted for various reasons (with political sensitivities being perhaps a major impetus), might have been predicted by the theoretical literature. The original articulation of yardstick competition by Shleifer (1985), for example, highlighted some of the potential difficulties that would need addressing for it to offer the optimum outcome. Whilst this form of performance regime reduces the costs of collecting information for the regulator, and thus can be a powerful tool in generating performance improvement, it creates a complex environment in which agents can collude and respond in unintended ways. In addition, the design of relevant measures and their reliability are major issues. For example, if units are heterogeneous, in terms of their characteristics, and the regulator does not take proper account of this, then the result is a very inefficient use of
information. Units may be classified as poor performers, signalling a need to aspire to the practices of the ‘best’, when in fact they may be performing adequately in other areas but be constrained by local circumstances. The incentive to invest in capacity for improving future performance locally may also be dampened in yardstick competition as such improvements are unobservable to the regulator (Chong, 2004). An evolutionary perspective also indicates that system changes result from the regulator learning from the behaviour, over time, of those whose performance is being managed (Courty and Marschke, 2003). These issues are apparent from the analysis of social care presented here and recent policies in England, in particular, may be attempting to resolve some of the limitations of this type of performance regime. However, it remains to be seen whether the alternative local performance measurement system, with reduced national targets, will offer the appropriate inducements for performance improvement by local agencies.
References


