

Just Do It? A Cautionary Tale on Implementing Performance Management Regimes

Joseph Drew

And

Sasindu Gamage

Abstract: *There has been a recent trend towards the use of performance indicators – often predicated on financial data – to enhance accountability and transparency of local government both in Australia and abroad. However, performance management also brings with it a range of risks including inter alia unintended performance distortion, synecdochal gap and intended distortions. This chapter reviews the substantial scholarly literature on government performance management before applying the concepts to a particular instance of high stakes performance management: the New South Wales Fit for the Future program. Fit for the Future required councils to self-assess against seven ratios drawn from financial statement data. Councils which failed to achieve the prescribed benchmarks were subject to forced amalgamations. Empirical evidence is provided which suggests significant levels of distortion in the performance management data. It is argued that careful design and testing of ratios – in order to avoid deleterious outcomes – is extremely important in any performance management regime, irrespective of whether accounting data from audited financial statements is used.*

Introduction

There is a vast scholarly literature on performance management in the context of the public sector. The literature is mostly pessimistic with respect to the efficacy of performance management and has a large focus on the performance paradox (the discord between performance on paper and actual performance), gaming (the exploiting of grey areas to present a more favourable impression of performance than might otherwise be warranted; Bevan and Hood 2006), and motivation (with particular reference to the dominant dichotomy of knights and knaves; Le Grand 2003). However, two rather curious features stand out in relation to the extant literature: First, the leading scholars in the field are drawn from the ranks of political science and management and second, (perhaps as a consequence), there is only a nascent literature on the use of audited financial data for compiling performance indicators. Yet, as we will demonstrate below, there are specific challenges posed as a result of employing audited financial data. To explicate on this matter we review the use of audited financial accrual data in the recent forced amalgamations of New South Wales (NSW) councils – dubbed *Fit for the Future*.

Fit for the Future (FFTF) is a particularly useful case study to employ for the purposes of explicating on the challenges faced in employing audited financial data in high stakes public performance management. This is because different incentives were offered, appealing to both sides of the public sector motivation dichotomy, at discrete periods in time. As a result we have a valuable natural experiment regarding the response of executives to increasing levels of incentives over time. FFTF was initially conceived as a sector led reform within the constraints of a firm state government policy of no forced amalgamation (Drew & Grant 2016a). However, in January 2014 the Independent Local Government Review Panel

(ILGRP) – charged by the state government to investigate evidenced based reforms to enhance the sustainability of local government in NSW – stunned the sector by releasing a report which called for ‘amalgamations [as] another essential component of reform, notably in metropolitan Sydney’ (Independent Local Government Review Panel 2013, p. 15). The major empirical evidence relied upon by the ILGRP was a sustainability assessment, based on ten financial ratios, conducted by the NSW Treasury Corporation (TCorp). Thus, well before the end of the 2014 financial year local governments were alerted to the fact that *amalgamations were proposed* and that financial data were being used to justify same. This was followed by the state government launch of a self-assessment programme in September 2014 whereby councils were directed to demonstrate financial fitness with reference to just seven (financial) performance indicators and propose ‘voluntary’ amalgamations (by 30th June 2015) in the event that they did not satisfy certain benchmarks. Thus, the stakes were raised from *mooted amalgamation* to an *amalgamation directive* by the end of the 2015 financial year.

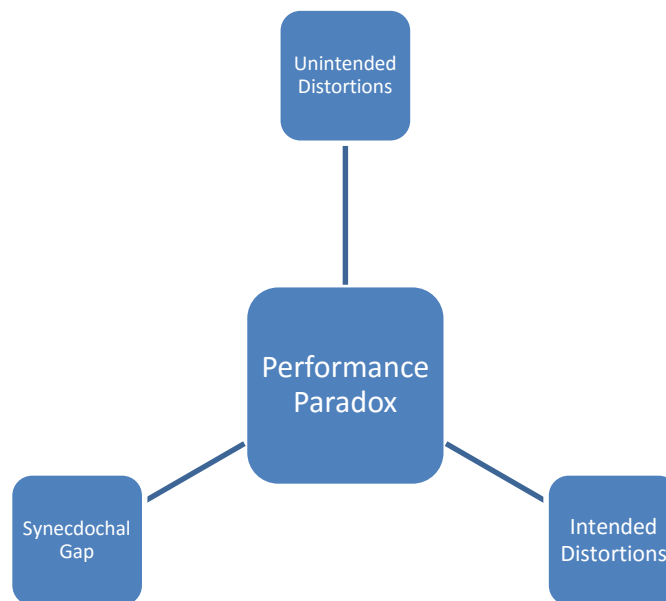
The aim of this chapter is to demonstrate that employing audited accounting data as the basis of a performance management regime in no way guarantees that the regime will be free of material distortion (and hence represent a sound basis for important public policy decision making).

The balance of the chapter is organised as follows. First, we review the extant literature on performance management with a particular emphasis on the unique challenges posed when audited financial data forms the foundation of performance indicators. We then set out the two fundamental assumptions which must be met in order for valid comparisons to be made with respect to financial performance indicators and outline our empirical methodology. This is followed by a discussion of our results which suggest that neither of the necessary assumptions were satisfied in the case of FFTF, therefore casting doubt on the empirical legitimacy of the entire programme. This chapter concludes with an enumeration of the important lessons for public policy architects intending to employ financial data for high stakes performance management.

Lessons from the Literature

The performance paradox is a discord between performance on paper and actual performance and is well attested in the corpus of scholarly literature on performance management (see, for example, Bevan and Hood (2006)). There are three causes of the performance paradox (see Figure 1) – unintended distortions (measurement error and the like), synecdochal gap (the disparity between the whole entity and the parts which are measured) and intended distortions (which might be either cheating (outright fabrication) or gaming (exploiting the grey areas)).

Figure 1 The Three Causes of Performance Paradox.



Unintended Distortions

Unintended distortion is undoubtedly the most neglected cause of the performance paradox, however, it has particular importance for performance management based on audited financial data – frequent users of financial report data are aware that misstatements do occur and that these occurrences persist despite the presence of internal and external auditing. A salient example of this is the recent revelation of ‘an \$8.68 million depreciation “system error”’ at Newcastle City Council (McCarthy 2014). No doubt misstatements also occur in non-audited financial data – the important difference is that in the case of accounting data there is a widespread confusion regarding the meaning of internal and external auditing of financial data. Many end users are under the erroneous impression that auditing is an absolute assurance that the data is free of error, when in fact it is only intended as *reasonable assurance* that the data is free of *material misstatement* (Drew & Dollery 2015). Moreover, what is considered material in a financial statement sense is very different to what might be considered material in a performance management sense – particularly if dichotomous benchmarks are detailed in advance (as was the case in FFTF). That is, if pass/fail benchmarks are detailed in advance and strictly observed, then materiality comes down to the

adjustment which is necessary in order to correct what might otherwise be a performance deficit for the entity.

Synecdochal Gap

When we come to the question of synecdochal gap we find a much more thorough treatment in the literature (see, for instance, Pollit (2013); Bevan and Hood (2006)). Synecdoche is an Aristotelian rhetorical trope in which a part is taken to represent a whole (*pars pro toto*) or the whole is taken for the part (*totum pro parte*) (Hamilton 2003). When it comes to performance management, focus has rested on the disparity between the whole (which is comprised of things both measurable and unmeasurable) and the part which is actually measured.

Synecdochal gap is problematic because naïve interpretation by end users can result in some seriously misleading conclusions (Walker, Boyne & Brewer 2010). This presents particular problems when financial data is used because lifting just a few pieces of information out of financial statements, which often run to a hundred or more pages, is not only an extreme case of synecdoche but also denudes the numbers used of the sense that they are, in fact, comprised of countless transactions and professional judgements. In so doing, performance indicators based on financial data may give an unwarranted impression of ontological truth (at the expense of a prudent cognisance of the role of professional accounting judgement). Moreover, the received wisdom is that the degree of synecdochal gap is proportional to the amount of gaming and strategic behaviour opportunities.

Intended Distortions

Gaming has attracted the bulk of the attention of scholars – not least because of the ‘thrill of the chase’ which one feels when one uncovers discrepancies which invalidate the claim of politicians seeking to employ performance management data for political purposes (see, for instance, Drew (2017); Drew and Grant (2016b); Bevan and Hood (2006); Drew and Dollery (2015b)). Before examining the types of gaming it is important to establish the difference between gaming and cheating. Gaming exploits grey areas to put the best gloss possible on performance, whereas cheating involves outright fabrication. In an accounting sense, gaming may be manipulation of discretionary accruals data to achieve a particular purpose (for example income smoothing – see, Copeland (1968); Marquardt and Wiedman (2004)), whilst cheating is fraud pure and simple (Rezaee 2005). In the former case, one can rely on professional judgement to get out of jail, in the latter there is no defence. Gaming has three main effects: threshold effects, ratchet effects and strategic behaviour. Threshold effects occur when performance benchmarks are communicated in advance – in this case there is clearly an incentive to perform to the benchmark level, but no incentive to perform at an entity’s utmost capacity. Ratchet effects, on the other hand, occur when executives rightly or wrongly perceive that performance in the present period will be used to set benchmarks for performance in later periods. If this belief is held then executives are actively dis-incentivised to perform above the bare minimum. The parallels between ratchet effects and the near zero balance public accounts literature (whereby it is hypothesised that executives manipulate discretionary accruals – particularly depreciation accruals – to achieve ‘Goldilocks’ results) are striking (see, for instance, Stalebrink (2007); Pilcher and Van der Zahn (2010)). By way

of contrast, strategic behaviour occurs when executives of entities emphasise only the areas of performance which are measured or pursue short term goals at the expense of long term performance. When solely financial data is used strategic behaviour may take the form of reducing service levels in order to enhance operating results (Drew & Dollery 2015) or extending the useful life of assets (knowing that there will be a cost when assets are ultimately sold or scrapped; Marquardt & Wiedman 2004).

The extent of intended distortions within a performance monitoring regime is largely dependent on the conjunction between dominant public sector motivations and the specific incentives posed to executives. In a well-known dichotomous framework Julian Le Grand (2010, p. 56) portrayed public servants as being either knights (motivated by ‘a spirit of altruism and the desire to perform a public service’ regardless of personal cost) or ‘knaves’ (motivated ‘completely by self-interest’). Bevan and Hood (2006) note that all performance management regimes must assume that the ranks are dominated by the former rather than the latter and, moreover, that the balance of knaves and knights does not shift as a result of incentives posed by the regime. In most instances, gaming and other subversive practices are attributed to knaves. However, Drew (2017) provides a convincing argument that there are situations in which knights may also participate in subversive activities. To understand this claim it is useful to think about the framework for accounting fraud incentive¹ outlined by Rezaee (2005). Rezaee (2005, p. 283) proposes just three incentives which explain most fraud – egocentric incentives (which clearly map nicely onto knavish motivations and the extant literature on gaming of public performance management regimes), economic pressure (which is not so relevant to public servants in the absence of performance bonuses and share prices) and ideological apologetics (which allow ‘individuals to think their behaviour or cause is morally superior’). Drew and Grant (2016b) contend that ideological apologetics can, in quite specific circumstances, result in knights also engaging in subversive activity. For instance, it is argued that local government amalgamations ‘have a huge impact on our community both in cost and in the loss of our community’s interest’ (Mayor Zelter of Woolahra council in Glanville 2016). If affected council executives were to believe claims of impending doom for their local communities than this might indeed result in a conjunction between knightly motivations and ideological apologetic incentives. Moreover, were this to be the case – that incentives were posed for both sides of the public service motivation dichotomy – then we might expect to see evidence of wide-spread gaming.

Specific Challenges Associated with Using Accounting Data

Thus, we can see that the use of accounting data as the basis for high stakes performance management regimes in no way escapes the well documented pitfalls detailed in the extant literature. In fact, use of accounting data also poses additional problems which have largely escaped the attention of the scholarly community (notable exceptions include Drew, 2017; Drew and Grant 2016b). First, there is an additional clear mechanism for gaming – which provides perpetrators with a ready excuse of ‘professional judgement’ (in the case of distortion of depreciation accruals). In FFTF three of the seven ratios could be manipulated by distorting depreciation accrual data – the operating ratio (operating surplus divided by operating revenue), the buildings and infrastructure renewal ratio (asset renewals divided by

depreciation on infrastructure assets) and the ‘efficiency’ ratio (operating expenditure divided by resident population) – and our (subsequent) analysis focuses on this accrual item. Second, the NSW state government fell into the trap of eschewing performance management data audits presumably on the basis that most of the data inputs came from audited financial data. Yet, as noted above, the purpose (and materiality concerns) of financial statement audits are quite different to the purpose and materiality concerns of performance management audits – particularly where the benchmarks are nominated in advance. The final additional problem posed by use of accounting data in performance management is that the performance indicators thus generated tend to obscure the accounting judgements which lie behind the numbers. For instance, it would not have been evident to most end users that nominal dollars per capita cited for the ‘efficiency’ ratios were dependent on depreciation accrual estimates (which represent around a fifth of NSW local government expenditure; Grant & Drew, 2016).

We now outline two crucial implicit assumptions which must be true for FFTF process to have been empirically legitimate.

Methodology

In order to make valid decisions (and implicit comparisons) on council financial performance indicators it is clearly necessary to accept two important assumptions:

- (i) That financial inputs (to performance indicators) are the result of consistent practice across the cohort; and
- (ii) That the egocentric and ideological apologetic incentives do not result in a material change to practice.

Our focus is on the depreciation accruals which are an obvious vehicle for material gaming as per our discussion in section 2 (although clearly the assumptions apply to all inputs).

Marquardt and Wiedman (2004), Pilcher and Van der Zahn (2010), Mulford and Comiskey (2002) and others have measured the quantum of unexpected depreciation according to the following algorithm:

$$UDEP_{j,t} = [DEP_{j,t} - (DEP_{j,t-1} * \text{Gross IPPE}_{j,t} / \text{Gross IPPE}_{j,t-1})]$$

Where UDEP is unexpected depreciation, DEP is depreciation, and IPPE is Infrastructure Property Plant and Equipment.

In simple terms, the methodology asserts that the rate of depreciation should remain consistent in consecutive financial years and can thus form the basis for calculating expected depreciation (and hence unexpected depreciation) with reference to current carrying values. We also employ this method – initially we adapt the general premise of comparing rates of depreciation in order to assess whether the first assumption was valid, before applying the method more directly to assess the second assumption.

In both cases we have elected to make a few improvements to the methodology in order to enhance its accuracy. First, it seems important to exclude non-depreciable IPPE assets in

view of the fact that around 37% of IPPE in NSW local government accounts is non-depreciable (Drew & Grant, 2016a). Failure to remove these items (a prominent example is land) confounds the analysis, particularly in view of the fact that individual items are often subject to contrariwise (upwards) revaluations. Therefore, UDEP and DEP in our analysis refer to the depreciable portion of IPPE only. Second, we believe that the unexpected depreciation should be calculated separately for each class of assets given that discordant useful lives (of say buildings, with lives of 80 years or more compared with books, with lives of 3 years or so) would otherwise result in conflated expected depreciation quantum (this is particularly important when new long-lived assets are acquired by a local government during the period of study). We therefore made 18 separate calculations (one for each distinct class of depreciable IPPE asset) and summed same in lieu of the one calculation employed in the extant literature. Notably, acceptance of the need for this second innovation to the methodology necessarily dictates that the first innovation is also observed. We also expressed the unexpected depreciation as the percentage variance from expected depreciation in order to facilitate valid comparisons across the cohort².

To test the first assumption we calculated the average rate of depreciation for each of the 18 distinct depreciable asset classes in 2013 over the cohort of 38 councils in the Greater Sydney region and used this average rate in place of 'DEP_{j,t-1}/Gross IPPE_{j,t-1}' in the above algorithm. We then expressed the sum of quantum of unexpected depreciation for each asset class as a percentage of expected depreciation for each council. The results were summarised (see, Table 1) with particular attention to measures of spread as a basis of assessing the validity of the reasonableness assumption.

To test the second assumption we used the rate of depreciation for each asset class in the 2013 financial year as the basis of calculating the quantum of expected depreciation in 2014 *and* 2015 respectively with reference to the relevant carrying value for each asset class. It should be noted that the rate of depreciation used was the specific rate employed by each individual council in the 2013 financial year, (not the average rate). We used the 2013 asset class rate of depreciation for both years because at the time that the 2013 financial year accounts were being finalised there was no suggestion that council amalgamations might occur. The 2013 rates of depreciation are thus free of the incentives for gaming which occurred in 2014 (when amalgamations were officially mooted) and 2015 (when councils were directed to pursue amalgamations). The results of these analyses are detailed in table 2 below.

Finally, to test the composite effect of both assumptions being violated we calculated the consistency of depreciation practice (as per the method employed for testing assumption 1) and present results of same in table 3.

Results

For valid comparisons to be made on the financial performance indicators it is clearly necessary to assume that practice is more or less consistent (certainly that the degree of inconsistency is immaterial with respect to achievement or non-achievement of the pre-

determined performance benchmarks). This was always a rather heroic assumption given extant scholarly evidence to the contrary (see, for example Pilcher (2006)) along with the reservations expressed by the Audit Office of NSW (2012)). Moreover, as the measure of spread (inter quartile range) in table 1 shows, this assumption was violated by a rather alarming margin. Clearly if practice was not consistent before the advent of FFTF, there is little reason to expect that fair comparisons could be made for the three performance indicator benchmarks which used depreciation data. We acknowledge that there might be some legitimate variation owing to, for instance, different classification of assets within the 18 different IPPE classes – however the scale of the variation suggests that there is something more at work here.

Table 1. Inconsistent Depreciation, Greater Sydney Councils 2013 (percent) – testing the first assumption

Year	Smallest	Largest	Q1	Median	Q3	IQR
Inconsistent Depreciation 2013	-39.21	45.27	-19.44	-4.085	25.3	44.74

Moreover, as Table 2 demonstrates the incentives posed by FFTF resulted in rather startling changes to practice in both the 2014 and 2015 financial years. Specifically, the typical council adjusted depreciation practice downwards (median of -2.83% in 2014 and -6.535% in 2015) in an apparent effort to convey a more favourable impression of performance against the FFTF criteria. Moreover, the downward shift seems to have become more pronounced as the incentives became greater (that is, as we moved from mooted amalgamation in 2014 to an amalgamation directive in 2015). In addition, the shifts in behaviour seem to have been widespread – thus suggesting that the particular incentives posed by FFTF found a response in a large proportion of the local government executive ranks. This observation is in turn consistent with our conjecture that there are circumstances (such as local government amalgamation) which might give rise to a conjunction of knightly *and* knavish motivations (which stands in contrast to the scholarly consensus that only knaves are responsible for subversive activity in performance management regimes – see, for example Bevan and Hood, 2006). All of this evidence is consistent with the media reports that that ‘some councils..... “played games” with their books by extending the useful life of their public assets to an “unrealistic” age’ (Bell 2015).

Table 2. Unexpected Depreciation, Greater Sydney Councils 2014-2015 (percent)

Year	Smallest	Largest	Q1	Median	Q3	IQR
Unexpected Depreciation 2014 - mooted amalgamation	-38.42	10.26	-11.67	-2.83	1.24	12.91
Unexpected Depreciation 2015 – amalgamation directive	-46.07	16.46	-20.19	-6.535	0.86	21.05

However, the question remains as to the overall consequence of inconsistent practice *combined* with subsequent changes to practice elicited from egocentric and ideological apologetic incentives. There are two ways of thinking about this composite effect. The first is to replicate the methodology employed to test the first assumption (but use 2015 data instead) – this approach tells us how much variation in practice existed with the cohort of local governments at the end of the FFTF process and is presented in the first row of table 3. Curiously the spread of results (as measured by the interquartile range) narrowed from 2013 levels, suggesting that there was some convergence in practice by the end of 2015. However, practice was still far too discordant for anyone to believe that valid judgements could be made regarding fitness for the future on the basis of the data. Moreover, whilst convergence in practice might have occurred we have no way of thinking about how the practice for the whole cohort had also changed in response to incentives posed by FFTF. Therefore in the second row we calculate the inconsistent depreciation practice across the cohort by employing the depreciation rates as at 2013 (before incentives for distortion became apparent). Once again, the results suggest a convergence in behaviour. However, the results also clearly show a shift downwards in the depreciation practice of the entire cohort (which might be expected given the results detailed in table 2). What this means is that there is strong empirical evidence that FFTF did significantly reduce depreciation accruals, notwithstanding the fact that there was some convergence amongst peers.

Table 3. Inconsistent Depreciation, Greater Sydney Councils 2015 (percent) – testing the composite effect of violating both assumptions.

Year	Smallest	Largest	Q1	Median	Q3	IQR
Inconsistent Depreciation 2015 (using 2015 depreciation rates)	-43.48	67.19	-24.22	-7.33	14.25	38.47
Inconsistent Depreciation 2015 (using 2013 depreciation rates)	-46.59	57.42	-26.67	-11.54	8.22	34.89

In sum, the empirical evidence suggests that audited financial data is not, in and of itself, a sufficiently firm foundation for high stakes performance management regimes. Moreover, we note that the ‘pass’ or ‘fail’ assessments for each of the three performance indicators (operating ratio, buildings and infrastructure renewal ratio and efficiency ratio) provided absolutely no indication that one of the key financial inputs to the said ratios (depreciation) was typically distorted by around 11.54%. This ‘obscurement’ of the accounting judgements which underlie financial performance indicators is a key area of concern moving forward.

Discussion

The forgoing analysis has demonstrated that public performance management regimes based on financial data are not only subject to the large number of problems identified in the extant scholarly work on performance management, but also to additional problems specific to the nature of accounting data. These additional problems include new opportunities for gaming

based on professional (accounting) judgement, the temptation to eschew performance auditing and rely instead on financial statement auditing, and the likelihood that end users might remain oblivious to the fact that performance indicators are strongly reliant on accounting judgement. Specifically we note that significant levels of manipulation of depreciation accrual data undermine the essential assumptions of the FFTF performance management and that these distortions occurred despite the fact that the financial statements had been audited. We also note that the FFTF pass/fail dichotomy – in the absence of analysis such as this – gives no indication of the performance indicator’s reliance on heavily distorted depreciation data. This fact highlights the danger that accounting judgements are obscured when performance indicators are collated (especially in the absence of sensitivity analysis and concomitant indicators for margin of error).

In sum, it is clear that high stakes public performance management regimes require higher audit effort (and that reliance cannot be made of financial audit alone). This is particularly the case where performance benchmarks have been communicated in advance (as was the case for FFTF) as ‘complete specification of targets and how performance will be measured almost invites reactive gaming by managers’ (Bevan & Hood 2006 p. 533; and our evidence indicates that reliance on financial auditing does little to change this assessment). Moreover, the failure to take measures to combat gaming and other subversive actions is likely in time to provide testament to Bird et al. (2005, p. 10) assertion that ‘to buy cheap methodology is to buy dear in the longer term if subsequent audit or independent critique [such as our analysis above] discovers problems with performance data which have been influential in public policy debate’.

Conclusions

Lest this chapter be interpreted as a gloomy assessment of public policy performance management regimes we now detail some of the measures which might have been taken to prevent the decision to spend ‘up to a billion dollars’ on local government reform in NSW being made on clearly unreliable empirical data (Drew & Dollery 2015).

First and foremost, a system of thorough performance management audit must form the foundation of any competent performance management regime with emphasis on materiality with respect to otherwise performance deficit. Analysis of the type conducted above would have allowed regulators to understand the reliability of the performance indicators used in public policymaking. Moreover, sensitivity analysis might have been conducted with a view to constructing confidence intervals for performance indicators. Second, indicators might have been constructed with the purpose of discouraging manipulation of obvious targets for gaming – for instance a depreciation rate presented as a proportion of depreciable assets would have put potential gamers on notice that their efforts would attract attention (see, Drew and Dollery (2015)). In addition, the responsible regulator might have made executives aware that attention would be paid to unexpected movements in accrual items and that consequences would be imposed on individuals caught distorting data (a ‘hang the admirals’ approach as discussed by Bevan and Hood (2006)). Fourth, efforts might have been made to increase the unpredictability of the performance management product (process on the other

hand should be clearly outlined in order to promote trust in the regime. Clearly there are some constraints imposed when using financial data (the financial year is fixed) however, unpredictability can be introduced in other ways – for instance, a larger set of performance indicators might have been specified, but attention focussed on the smaller subset (although not so small as to elicit conclusions of extreme synecdoche). Another effective way of increasing the cost of gaming is to commit to continuing use of the performance indicators – in so doing, gamers are presented with a smaller subset of suitable gaming devices (changes which commit the organisation to future actions are no longer appealing) and are also more likely to incur the costs of gaming (for example, the cost of extending the useful life of assets will continue into a future time). In similar vein, mandating independent audit committees³, requiring audit committees to sign off on the performance management inputs and committing to a future review of FFTF outcomes would have increased the likelihood of gamers being detected.

In short, the emphasis should be on increasing the cost of gaming and ensuring end users are aware of the potential error imputed into individual performance indicators.

In view of both the unreliability of the performance indicators employed to guide the billion dollar reform to NSW local government and the fact that the embarrassing flaws might have been avoided, it is clear that the accounting profession needs to engage more fully in the development and scholarly evaluation of performance management regimes. Political and management scientists have made very important advances in many fields – but when it comes to performance management on the basis of financial data there is a clear role for the expertise which only an accounting professional can provide.

NOTES

¹ Although, as noted earlier, the focus of our chapter is on gaming rather than cheating the framework of incentives is still useful for understanding how and why the dichotomy of public sector motivations might respond to particular incentives.

² We concede that the methodological innovations (along with extant methodology) would fail to detect some manipulations such as incorrectly classifying assets as non-depreciable items or keeping items as Work in Progress (WIP) beyond commissioning date.

³ Only around half of NSW councils had audit committees at the commencement of FFTF. Moreover, most audit committees are not truly independent.

References

- Audit Office of New South Wales (AONSW) 2012, *New South Wales Auditor-General's report performance audit – monitoring local government*, AONSW, Sydney.
- Bell, A 2015, 'Blacktown Mayor Stephen Bali rejects IPART Fit for the Future findings', *Daily Telegraph*, 20 October, viewed 23 December 2016 2015, <http://www.dailytelegraph.com.au/newslocal/news/blacktown-mayor-stephen-bali-rejects-ipart-fit-for-the-future-findings/news-story/af9a5472ebb39fbcd43ebc11d0de58c>
- Bevan, G & Hood, C 2006, 'What's measured is what matters: targets and gaming in the English public health care system', *Public Administration*, vol. 84, no. 3, pp. 517-538.
- Bird, S, Cox, D, Farewell, V, Goldstein, H, Holt, T & Smith, P 2005, 'Performance indicators: good, bad, and ugly', *Journal of the Royal Statistical Academy*, vol. 168, no. 1, pp. 1-27.
- Copeland, R 1968, 'Income smoothing', *Journal of Accounting Research*, vol. 6, no. 1, pp. 101-116.
- De Bruijn, H 2007, *Managing performance in the public sector*, Routledge Taylor and Francis Group, London.
- De Bruijn, H & Van Helden, G 2006, 'A plea for dialogue driven performance-based management systems: evidence from the Dutch public sector', *Financial Accountability & Management*, vol. 22, no. 4, pp. 405-423.
- Drew, J 2017, 'Playing for keeps: local government distortion of depreciation accruals in response to high stakes public policy-making', *Public Money & Management*, In Print.
- Drew, J & Dollery, B 2015, 'Less haste more speed: the Fit for Future reform program in New South Wales local government', *Australian Journal of Public Administration*, vol. 75, no. 1, pp. 78-88.
- Drew, J & Grant, B 2016a, 'Multiple agents, blame games and public policy-making: the case of local government reform in New South Wales', *Australian Journal of Political Science*, doi: 10.1080/10361146.2016.1238872
- Drew, J & Grant, B 2016b, 'Means, motive and opportunity: distortion of public policy making performance management data', *Australian Journal of Public Administration*, doi: 10.1111/1467-8500.12225
- Drew, J, Grant, B & Campbell, N 2016, 'Progressive and reactionary rhetoric in the municipal reform debate in New South Wales (NSW) Australia', *Australian Journal of Political Science*, doi: 10.1080/10361146.2016.1154926
- Glanville, B 2016, 'Woollahra Council loses legal challenge against NSW Government over forced amalgamation', *Australian Broadcasting Commission*, 20 July, viewed 22 December

2016, <http://www.abc.net.au/news/2016-07-20/woollahra-council-loses-legal-challenge-over-forced-amalgamation/7643776>

Grant, B & Drew, J 2016, *Local government in Australia: history, theory and public policy*, Springer Palgrave, Singapore.

Hamilton, P 2003, 'The saliency of synecdoche: the part and the whole of employment relations', *Journal of Management Studies*, vol. 40, no. 7, pp. 1569-1585.

Hopwood, A 1987, 'The archaeology of accounting systems', *Accounting, Organisations and Society*, vol. 112, no. 3, pp. 207-234.

Independent Local Government Review Panel (ILGRP) 2013, *Revitalising local government: final report of the NSW Independent Local Government Review Panel*, ILGRP, Sydney.

Le Grand, J 2003, *Motivation, agency, and public policy: of knights & knaves, pawns & queens*, Oxford University Press, Oxford.

Le Grand, J 2010, 'Knights and knaves return: public service motivation and the delivery of public services', *International Public Management Journal*, vol. 13, no. 1, pp. 56-71.

Marquardt, C & Wiedman, C 2004, 'How are earnings managed? An examination of specific accruals', *Contemporary Accounting Research*, vol. 21, no. 2, pp. 461-491.

McCarthy, J 2014, 'Silence greets call to double check audits', *The Newcastle Herald*, 23 July, viewed 24 December 2014, <http://newsstore.smh.com.au/apps/viewDocument.ac?page=1&sy=afr&kw=Silence+greet+all+to+double+check+audits&pb=nch&dt=selectRange&dr=5years&so=relevance&sf=text&sf=headline&rc=10&rm=200&sp=nrm&clsPage=1&docID=NCH140723A46R235R4A4>

Pilcher, R 2006, 'The smoothing potential of depreciation for local authorities', *Journal of Contemporary Issues in Business and Government*, vol. 12, no. 2, pp. 67-80.

Pilcher, R & Van der Zahn, M 2010, 'Local governments, unexpected depreciation and financial performance adjustment', *Financial Accountability & Management*, vol. 26, no. 3, pp. 299-323.

Pollitt, C 2013, 'The logics of performance management', *Evaluation*, vol. 19, no. 4, pp. 346-363.

Rezaee, Z 2005, 'Causes, consequences, and deterrence of financial statement fraud', *Critical Perspectives on Accounting*, vol. 16, no. 1, pp. 227-298.

Stalebrink, O 2007, 'An investigation of discretionary accruals and surplus-deficit management: evidence from Swedish municipalities', *Financial Accountability & Management*, vol. 23, no. 4, pp. 441-458.

TCorp 2013, *Financial sustainability of the New South Wales local government sector*, TCorp, Sydney.

Walker, R, Boyne, G & Brewer, G (eds.) 2010, *Public management and performance: research directions*, Cambridge University Press, Cambridge