CARR Regulators' Forum: Third Meeting

Regulatory Performance

There is said to be a growing pressure on regulators to account for and justify their performance. Questions about how to manage and measure performance link to three central concerns. One is the purpose of performance management; different rationales can be distinguished, ranging from those interested in financial and overall outcomes, to those interested in encouraging learning, to those seeking to motivate and celebrate achievements of the industry or of the regulatory agency itself. A second, related concern is the many ways in which performance can be measured, and where the methods of performance measurement are not necessarily aligned to the overall objective. Third, performance management impacts on behaviours, raising concerns about distortions, such as gaming behaviour. These questions affect all regulators.

Why, how and for whom is performance being measured?

Across regulators, there was agreement that there was some movement from measuring inputs to a more general concern with the effectiveness of interventions. These shifts had partly been in response to incidents and public criticism. Performance measures had been developed on the basis of theory-driven expectations of causal relationships. Results framed discussions about evaluation, which were conducted on the basis of a scorecard. Such exercises had two key values, one was to fulfil the need for enhanced accountability, namely the need to show that the regulatory activities were providing for 'value for money'. In addition, there was also a strategic value, which offered the wider public information to influence its choices, and allowed regulators to tackle poor performance in the regulated sector.

Nevertheless, there was agreement that there was also a degree of ambiguity. One could never be entirely certain as to whether a particular intervention had led to a specific improvement. There was often no clear causal model, there were many unknowns. Instead, it may be good to select a particular outcome and then try to explore how a particular intervention had contributed.

Numerous audiences for performance management could be identified. Subtle differences existed between all of these. Relevant audiences were the top management of the regulatory agencies themselves, the regulated industry, politicians and the wider public. It was at times difficult to address all these at the same time. Management needed to know that the organisation was achieving the things it should be doing. However, other stakeholders may have different demands. Deciding on which demand mattered most was a challenge. Potentially it also meant that there was pressure on the commitment to measure outcomes, because it was difficult to identify what 'good' actually was. Indeed, the argument in some areas was more about 'good enough', and there was uncertainty and ambiguity in moving beyond that level, especially if compliance data was not necessarily reliable. The art was to be able to be nimble and to please the diverse audiences.

For others, the question was about identifying failures in the market, and how the regulator was influencing behaviours. When looking at 'value for money', it was easier to look at economy and efficiency, but identifying effectiveness was more difficult. There was an issue of what could be measured, and non-events were very difficult to measure. Others noted that there was at times an emphasis on numbers rather than attention on overall effectiveness. Regulators noted that questions about effectiveness always proved challenging. While numbers were helpful, there should be an emphasis on trends over time; a measure of effectiveness should involve actual engagement with people. More generally, there was a problem when poor performance revealed an inherently dysfunctional organisation. It was questionable whether regulators could provide solutions to far wider-reaching problems in an industry. Leadership in the regulated industry seemed critical.

The key challenge for regulators was however to improve performance of the regulated industry beyond the minimum level, and the identification of 'bad apples'. The demand was to shift the whole industry's performance and this was a major challenge. Measuring minimum levels of compliance was possible, any further ambition was likely to run into difficulties of measurement.

This raised a further question about where to draw the line for regulatory attention. There were questions as to what outstanding behaviour looked like, and what kind of financial implications a move towards 'outstanding' would have for the regulated industry. In particular during times of an economic downturn, such a demand could have considerable resource implications for the industry.

More generally, the cost of regulation was increasingly becoming a concern. One had to show industry that the costs that regulators imposed were effective and reasonable. This raised again the question about whether one was measuring enough aspects to demonstrate progress. Nevertheless, one has to ask whether the industry would not eventually resist increased licence fees.

What kinds of performance management can be identified?

Across regulators, a range of measures were used, based on 'hard' and 'soft' data from internal and external information. Some regulators developed performance management systems in liaison with the regulated industry. Much of the data that was being used relied on compliance data. Some regulators could rely on a large over-time dataset, others were not able to rely on similarly extensive information. The data could be used to provide graphic illustrations, and also to rank performances. Such performance ranking was usually based on an assessment of the performance, as well as the responsiveness of the regulated industry to previous events.

Some regulators were overseeing other regulatory bodies. This meant that one could engage with industry associations and trade bodies. These would sometimes suggest that regulators should be tougher. One could then exert pressure on these regulatory bodies. For example, this could involve the publication of progress over certain time periods. Such 'naming and shaming' did lead to a positive response in terms of compliance. Some regulators noted how the publication of information had led to a private sector response, which in itself could encourage the addressing of regulatory problems. However, it was important to realise how reliable the information regarding compliance records actually was. One potential problem was how to understand why measured performances improved, and what the time-lag between the measurement and the receipt of the information was. Some regulators had effectively managed to receive more up-to-date information and reduced the problem of time-lag.

Nor was measuring performance straightforward, as 'improved' performance in terms of output measures was not necessarily a result of better performance of the regulated industry. Output measures were therefore not always appropriate when it came to the assessment of efficiency and effectiveness of oversight. Furthermore, as many regulatees were subject to the activities of different regulators, it was not always easy to trace any one regulator's impact on achieving altered behaviours.

How can one ensure that distortion effects do not emerge?

There were examples where regulators had realised that the way they measured performance had led to distorting effects. Just measuring the number of inspections was not seen as a good measure for encouraging improvements, Organisations always responded to targets. There were also issues when industry played off different regulatory regimes. It was important to be aware of perverse behaviours that may be induced by the activities of other regulatory organisations.

It was essential to move towards a model that was focused on outcomes. Examples existed where, with a shift in emphasis, a fall in the number of inspections had accompanied improved industry performance. Similarly, it was noted how a move from 'prohibition' to 'improvement' notifications had encouraged better performance, as it signalled that regulators were engaging differently with industry. Broadly, measuring inputs was seen as unhelpful for improving performance. A move towards outcomes was nevertheless seen as very difficult. It raised challenges in terms of understanding and learning from the data, how the data reflected on the targets, how not to come to rash decisions about the need for any changes, and being willing to disclose uncomfortable news. One had to be clear about why one was collecting information, and distinguish between purposes. The demand for formalising things was potentially problematic, as it took away any 'safe space'. Once things were measured formally, people might be frightened off from reporting problems. One challenge was to deal with the right direction of travel, and to ensure that there was consistency in inspection activities. Another was to ensure that not just good news was being celebrated. It was important to celebrate success, but also to encouraging learning from problems.