For my internship I prepared a short paper on social sector productivity for the Productivity Commission. The objective was to consider the implications of introducing productivity measurement in the social sector, given some of the complexities of observing outputs and outcomes for certain sector tasks. To do this I selected one typology, James Q. Wilson’s matrix of government tasks in Gregory (1995b), and attempted to apply this to a set of tasks within one organisation in the social sector, namely the Ministry of Social Development (MSD). The tasks were drawn from MSD’s annual report for 2015/16.

**Productivity in the social sector**
The social sector is complex, covering a variety of activities, including health, education and welfare services. Productivity measures the capacity of an economy, industry or organisation to produce goods and services (outputs) using inputs such as labour and capital (such as machinery, computer software and land). It is a quantitative measure of the ratio of the volume of output to the volume of inputs (Gemmill, Nolan and Scobie, 2017). Productivity is a useful concept in the social sector because delivering more or improved services with the same inputs (or the same services with fewer inputs) can potentially enhance well-being, all else being equal (Conway, 2016).

An important component of productivity measurement is quality adjustment, which ensures that quality changes are considered when measuring productivity, to ensure a fair picture of performance and minimise the promotion of productivity improvements at the expense of quality (Hanushek and Ettema, 2015). Consequently, to be able to apply standard productivity measures to social sector tasks we must be able to observe some defined input, and the output or outcome, and we must be able to quality-adjust the result.

**Theory**
Wilson’s typology differentiates among types of tasks by the observability of outputs and outcomes, and identifies four categories, as shown in Figure 1. Production tasks have observable outputs and outcomes; procedural ones have observable outputs but outcomes that are difficult to observe; craft ones have observable outcomes but outputs are difficult to observe; and it is difficult to observe both outputs and outcomes of coping tasks (Gregory, 1995a).

Wilson’s typology, and this analysis, employ a broad definition of outputs as the work that organisations carry out and the things (goods and services) produced, and outcomes as the effects of this work...
on communities and society at large (Gregory and Lonti, 2008). However, outputs and outcomes can be further broken down according to the level at which we wish to observe them, a point which will be returned to later.

Practice

Performance measurement practice in the social sector is governed by the requirements of the Public Finance Act 1989, which, in contrast to Wilson's typology, considers all organisations to be based on production-type tasks, with observable outputs and outcomes (Gregory, 1995b; Treasury, 2005). For many years the government has tried to shift performance measurement in the social sector to an outcomes focus (Destremau and Wilson, 2016, p.33). However, it can be difficult to observe (and, consequently, to measure) some outcomes, particularly those that only become evident in the long term (Alford 1993; Productivity Commission, 2015). Further, co-production can cause issues for measuring some social sector tasks. Co-production entails the contribution of people or organisations external to the producing organisation (such as the target group being regulated, or other public actors) to accomplish objectives (Alford, 1993). Co-production is usually required for more complex tasks, namely craft and coping tasks, and can cause significant issues for attributing outcomes to a single organisation by making it hard to unpick the causes and effects of contributing efforts (Gregory, 1995b). As craft tasks have some observable outcomes, the effects of co-production are slightly less significant for measurement, as there is some observable end which can be linked with production technologies (technical knowledge). However, in the case of coping tasks this matter can be challenging (ibid.). Nevertheless, much measurement already exists in the social sector, including for tasks which have complex, long-term goals, such as social work (see Ministry of Social Development, 2016).

Key findings

Limits of the typology

There are some clear limitations to Wilson’s typology when applying it to practice. Indeed, Wilson himself once warned that the typology should be used ‘with caution’ (1989, p.159). Significantly, the attempt to categorise many complex tasks into a small number of categories – in this case four – is unduly restrictive (Lonti and Gregory, 2007). The exercise of applying Wilson’s approach to the services undertaken by MSD (see Figures 2 and 3) illustrated this, and the lines between quadrants in Figure 2 have been dashed to represent that the classifications are not definitive.

In some instances, a classification that would allow a task to straddle the lines between quadrants might be most appropriate. For example, certain tasks can have sub-tasks that are more or less observable. For instance, fraud prevention is a sub-task of administering income support to seniors (Ministry of Social Development, 2015). Fraud prevention requires more discretion on the part of officials and thus is more difficult to observe and might not fit within the production quadrant. Nonetheless, other tasks fit relatively neatly within the typology, with care and protection services (social work) a clear coping task.

One-size-fits-all approach not sufficient

Despite the limitations, applying the typology to MSD highlighted that much task diversity exists within a single social sector organisation, as shown in Figure 2. This indicates that a one-size-fits-all approach would be inappropriate for productivity measures in the sector. Inevitably, the measurement of a coping task like social work will require a different approach from that required by a production task like processing payments (Gregory, 1995a). In these more complex cases, moving away from standard productivity measures and towards productivity-type measures, such as cost-effectiveness (i.e. measuring the relationship between inputs and final
Defining levels of measurement

The way we define outputs and outcomes affects how we might go about measuring them. Wilson's typology employs a broad definition of outputs and outcomes, but these can be further broken down according to the level at which we wish to observe them. Outputs can be observed either at the level of the specific day-to-day activities of individuals (as Wilson suggests), or at a higher-level overview of this work, such as the number of hours worked or the number of children seen in the case of a social worker (see Laking, 2008). Similarly, outcomes can be broken down into intermediate outcomes and ultimate outcomes, the former being more observable, shorter-term goals and the latter being the final desired effect of the task, which could take years to eventuate and is more consistent with Wilson's definitions (Coglianese 2012; Gregory 1995a).

There is a need for greater clarity about what we mean when we refer to outputs and outcomes, as using the same words interchangeably can create confusion, not least when exploring performance measurement issues. Certainly, it will be easier to observe a higher-level overview of an output, and an intermediate outcome, than it would be to observe outputs and outcomes according to Wilson's definitions, and the former set of definitions are more consistent with those employed in the Public Finance Act 1989. Nevertheless, which level of measurement is appropriate depends on the objective – what it is that we are trying to achieve or learn by measuring the outputs and outcomes of a certain task – as different goals will require different levels of detail. Approaching measurement in this way will help to avoid ‘hitting the target and missing the point’ (Bevan and Hood, 2006, p.421).

Conclusions

Wilson’s typology is a useful tool that raises interesting questions about social sector productivity measurement. In many cases standard productivity concepts are compatible with social sector tasks. In others, however, we may need to apply more innovative, bespoke methods to capture performance. Above all, it is imperative to define outputs and outcomes clearly and avoid a one-size-fitting-all approach.

Table 1 Compatibility of selected tasks with productivity concepts

<table>
<thead>
<tr>
<th>Task type</th>
<th>Inputs (labour and capital)</th>
<th>Outputs</th>
<th>Outcomes</th>
<th>Ability to measure quality change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production – Administering income support seniors</td>
<td>Capital (money, computers, buildings etc.)</td>
<td>Monetary transfers</td>
<td>Ability of seniors to maintain independence and social participation</td>
<td>Using outputs or outcomes</td>
</tr>
<tr>
<td></td>
<td>Labour intensive</td>
<td>Entitlement eligibility assessments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural – Data, analytics, evidence and policy advice</td>
<td>Capital (money, computers, buildings etc.)</td>
<td>Advice delivered to Minister</td>
<td>Unobservable – outcome attribution issues, impacts of work uncertain</td>
<td>Using outputs</td>
</tr>
<tr>
<td></td>
<td>Labour intensive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craft – Improving employment and social outcomes support</td>
<td>Capital (money, computers, buildings etc.)</td>
<td>Unobservable – much discretion by officials, difficult to prescribe outputs</td>
<td>Clients moving closer to independence (away from benefit dependency)</td>
<td>Using outcomes</td>
</tr>
<tr>
<td></td>
<td>Labour intensive</td>
<td>Attributes of clients etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coping – Care and protection services</td>
<td>Capital (money, computers, buildings etc.)</td>
<td>Unobservable – much discretion by officials, difficult to prescribe outputs</td>
<td>Unobservable – attribution problems, impacts of work uncertain</td>
<td>Difficult</td>
</tr>
<tr>
<td></td>
<td>Labour intensive</td>
<td>Attributes of clients etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3 The reliance of tasks on co-production

Source: Gregory, 1995a, p.174

New Zealand Productivity Commission Internship: social sector productivity
fits-all approach. Plainly, the social sector is too complex to be fully captured by Wilson’s typology, which consequently, as he warned, must be used ‘with caution’.

Acknowledgement
My internship was a great experience. My thanks to the Productivity Commission for the opportunity, and especially to Dr Patrick Nolan for his time in supervising my work. The views expressed here are my own.

References