3

PUBLIC SECTOR PRODUCTIVITY: PUZZLES, CONUNDRUMS, DILEMMAS AND THEIR SOLUTIONS

Patrick Dunleavy
London School of Economics

This chapter will discuss four aspects of productivity in the public sector. First, why do we have so little public sector productivity information? After all, productivity has been an important concept in the private sector for approximately 80 years, and has been much talked about since the early 1970s when President Nixon set up a Productivity Council in the US. However, what we have known about it up until now has been relatively restricted.

I want to try and explain why, on the whole, a lot of people in society— mostly economists and the right-wing press—tend to think of productivity as being completely flat in the public sector, and sometimes even declining. I will argue that we have no evidence for that. In fact, I will show that public sector productivity can grow spectacularly—and quickly—in some cases; and it can be flat, or even declining, in other cases. We need to look at this in a consistent way, and I will use three narratives running through some very big UK departments to tell these different stories.

Then I will pull together content from my book with Leandro Carrera (Dunleavy and Carrera 2013), as well as the academic literature on factors that shape government productivity. Many people have faith-based views of this—they say that the private sector is more productive than the public sector, and consequently they assume that we need more contestability and competition. There is zero evidence that competition and contestability make any difference to the public sector's productivity, for very good and strong reasons. I believe that rather than making faith-based decisions on how we set up the public services, we should look at the empirical evidence.

Finally, to round off our discussion, I will look at eight steps that our book suggests are important; four of them are rather immediate, and four are more long-term.

Productivity and the public sector—what is 'productivity'?

I am a conservative minimalist when it comes to how we define public sector productivity. It is a really simple concept and we should keep it that way: productivity is outputs divided by inputs. *Outputs* are not the same as *outcomes*—they do *not* include all the things we are trying to achieve in life, or the social transformation or public value that we may be trying to achieve, but rather they are the things that are relatively easy to measure and under our own control. An output would be a teacher delivering a lesson, or a fire service answering an emergency call. We could limit them to activities, but fundamentally outputs are things that an organisation provides and controls.

To measure productivity, we do not have to cover all of these outputs; it is usually sufficient to look at just a few of them. For example, for our research we looked at the productivity of the tax agency in the UK, Her Majesty's Revenue and Customs (HMRC). One of the biggest agencies in the UK government, it comprises a quarter of the staff in the entire UK civil service, and we needed to measure 14 different indicators to understand HMRC's productivity. In contrast, in studying most organisations we can obtain sufficient information by examining just two or three indicators of primary outputs or kinds of activities.

The next issue is measuring the value of outputs. In the private sector, when we have a firm that is producing lots of widgets, and an industry that is also producing lots of widgets, we can see how much they sell for, and we can weigh the overall output of the organisation by the prices that they sell for in the market. That is a very easy way of determining an overall level of output for the firm.

We cannot do that in the public sector, where we have a sizable problem determining the value of outputs. If our private sector firm produces a 'super widget' compared to last year's widget, and then sells it at a higher price, we can draw the welfare implication that somebody felt it was worth paying that extra price (or they would not have bought it). This means that in the private sector we can cope with technological and other types of change, and we can think about the welfare implications more easily. We cannot do any of that in the public sector, and consequently for the best part of the last 80 years, economists and national statisticians have treated government productivity as completely flat. They have reasoned that we cannot measure outputs, so instead of outputs we use inputs, and we divide inputs by inputs—with predictable results. This sounds very crude, but that is the way national statistics have been done and are still being done over most of the world.

Some progress on this matter can be attributed to economist Sir Tony Atkinson, who in a nutshell argued that we can get to grips with the overall output of a public sector organisation by looking at the things that it does and weighting them in relation to administrative costs rather than prices (Atkinson 2005). For example, if it is 10 times more expensive to collect Tax A than Tax B, then Tax A must count more in the overall output weight of the organisation.

Another problem in the public sector is that when somebody says 'productivity', everybody says, 'Yes, that's a great buzzword, I like that. Productivity means basically everything good.' So we tend to take this useful concept and inflate it. Productivity is a small concept that is part and parcel of what every organisation is doing—but it is not the whole story. If anybody says something like, 'We're looking at the social productivity of an organisation', it does not actually mean anything at all.

Issues in measuring public sector productivity

The next set of problems with public sector productivity relate to the fact that it is difficult to measure. It would be ideal to measure public sector productivity at a local level, or at a state level in a big federation like the USA, and so compare different sorts of units; hopefully the units would be doing pretty much the same sorts of things. That kind of data would be very useful. Unfortunately, if something has been done at the local level, or even at the state level, it is usually because it is not easy to standardise—it may be a personalised service, or needs to be delivered directly to people, because quality makes an enormous difference to whether the service is delivered correctly.

We could, for example, easily get perverse effects. Imagine a patient checks into Hospital A, and is treated very well. He/she is kept in hospital for just the right amount of time (e.g. four days), is released upon improvement and recovers. That counts as one unit of output. On the other hand, another patient suffering a similar complaint checks into Hospital B and is treated badly, discharged after only two days, and needs to be readmitted later on. Hospital B can then claim two units of output, even though they are doing a worse job than Hospital A. So we should be aware of possible perverse effects, and should also think about what we are trying to do. If one fire service answers a lot of fire calls, and another answers fewer fire calls because it does better fire prevention work, it is easy to misread the data and get the wrong signals. As a result, although it would be best to study productivity comparatively, it is often quite difficult to do.

At the national level, we often have just one agency—one defence agency, one tax agency, one social security agency, and so on. It is consequently hard to make comparisons, and the international data on productivity is substandard, despite years of work by the UN and OECD and other bodies. We do not have big N data sets; we have this single organisation. So the question is how can we measure productivity here?

It is also difficult to think about productivity patterns over time. Sometimes productivity will go up or down for reasons that are not under our control. If demand goes up, most organisations can cope with that for a year or two by cramming in more people and working the staff hard. But this is not really sustainable. Sometimes demand goes down—for example, we might roster extra staff to assess passports, but if fewer than expected foreign visitors come in then productivity drops before we can de-roster some of our staff.

There tends to be a very high level of resistance towards measuring productivity by politicians and civil servants at the national level or the state level. People dream up lots of excuses about why we should avoid measurement, many of which are based on the idea that the public service is really complicated. But I disagree—I do not think that public sector organisations are more complicated than private sector organisations.

As you can see, there are a lot of problems with measuring productivity, which is another reason not to get into things like outcomes. Keep it simple—it is hard enough to assess productivity correctly in the first place without bringing in extraneous elements.

Three case studies from the UK

In our five-year study, Leandro Carrera looked at six big agencies in Britain. This chapter will discuss three of them. I will focus primarily on some of the 'over time' evolution of productivity: if we want to know how an organisation is doing, and we can assemble some data that goes back even five years, we will begin to see patterns. We can then begin to think about why certain things are moving in particular ways, and why the ratio of outputs divided by inputs changes.

In our study, we focused on the big departments, but also ones with fairly standard outputs so that we had access to consistent, reliable data. We did not really look at effectiveness, or at the issue of how a department responds to a new government. Instead, we examined issues like how many tax forms were returned, how many staff there were, and so on—we kept it fairly simple. We also did not want to introduce quality changes unless necessary. For example, at the start of the period we were studying, only about 10 per cent of people were filling in their tax forms online. Now it is 84 per cent, and public servants at HMRC argued that this increase should count as a quality improvement. We decided that this does not count as a quality improvement because the rest of society has changed—filling in a 50-page paper form is no longer what anybody is doing anywhere else in society. Everybody is doing it online, so there is no extra quality credit for doing the same thing that everybody else is doing, and somewhat behind the pack at that.

What we need to do then is look at a detailed narrative for each agency. We suggest agency leaders should do this, because they know a lot about their own agencies, and where all the outputs and inputs are—they will know how many people are on staff, how much capital has been invested, the number of intermediate outputs they are getting from contractors, and so on. From this data, we can create an overall measure of output, weigh that output by the cost of doing different things, and so start to obtain an overall measure of productivity by dividing these figures by our inputs.

I will use three case studies to illustrate how these techniques can be useful:

- British Customs;
- HMRC, our new integrated tax agency created in 2005; and,
- Department of Work and Pensions (WAP), which handles social security.

HMRC and WAP between them account for half of the British civil service.
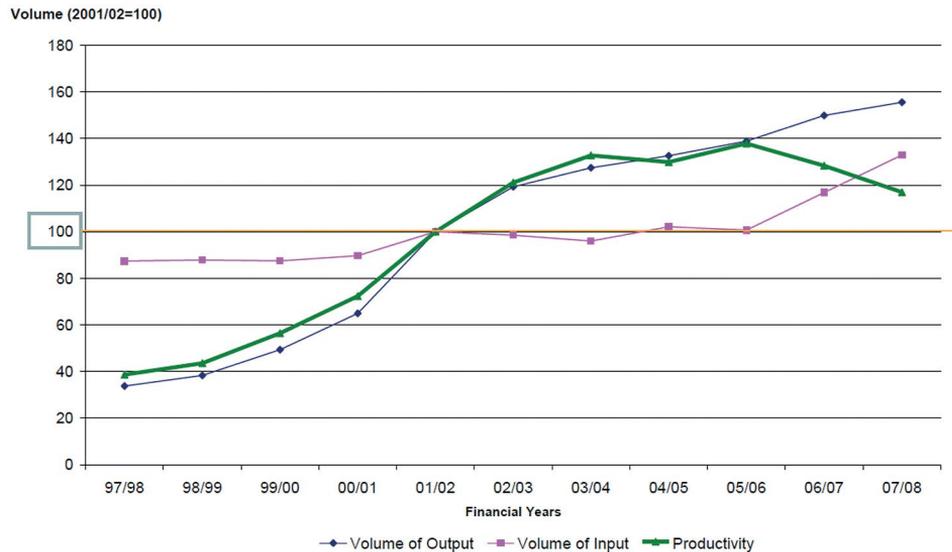
British Customs

Volume (2001/02=100)

Figure 3.1. Total factor productivity in the UK Customs regulation on trade, 1997 to 2008

Source: Patrick Dunleavy.

Figure 3.1 shows Customs regulating imports and exports. The volume of output is shown by the dark blue line, the volume of input is shown by the pink line, and the overall productivity index is shown by the thick green line. Over this period there was not much staff growth until 2005—so for 10 years the staff was constant. Meanwhile, the amount of containers coming into the UK, that is the amount of trade to be looked at (the green line), went up significantly because of the boom years before the Global Financial Crisis. When outputs are going up but there is no increase in inputs, productivity rises—and it rises very dramatically. If you looked at the USA you would see the same pattern in US Customs.

How did that happen? Firstly, Customs made a very big decision in the mid-1990s to introduce a computerised system. A pre-internet system for tracking all imports and exports, where everything was done online, this allowed Customs to make huge steps forward throughout the next 10 years without increasing their staff numbers—they were simply able to get more done. Of course, technology on its own does not really help unless you make other kinds of changes. British Customs used to operate on a volumetric basis; the general practice was to open up every 20th or 50th container to see what was in it. The containers were opened blind, and sometimes problems were found, and sometimes they were not. It was not a very sophisticated way of doing things.

Over time, British Customs more than any other customs organisation in the world transformed its processes into a risk-based and intelligence-based way of looking at import and export trade—despite the fact that they open far fewer containers now than any other country in the European Union.

One reason they managed to achieve this huge change is the technological investment discussed earlier, and the other is the managerial changes that made a significantly more effective use of the same number of people. Even here we can see in Figure 3.1 that towards the end of the period, there is a bit of an increase in inputs, and productivity begins to tail off a little. But the overall change is very dramatic—any private sector organisation would be very pleased to achieve the same threefold increase in productivity over time.

Her Majesty's Revenue and Customs
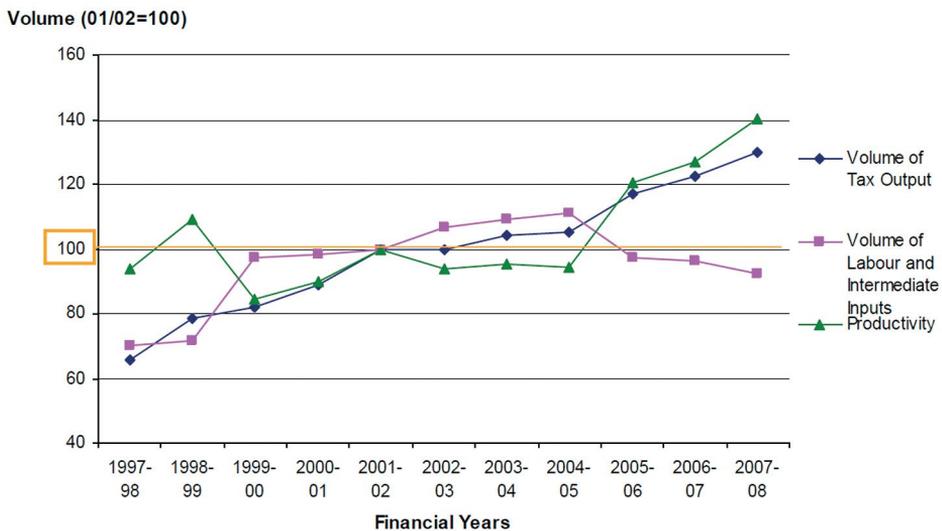
Volume (01/02=100)

Figure 3.2. Labour and intermediate inputs productivity, UK taxation, 1997 to 2008, using tax collection activity data

Source: Patrick Dunleavy.

Unfortunately, we do not have complete data on the capital investments for HMRC, so we used a measure looking only at staff and contractors—there is no capital investment reflected in the charts. While this is a bit of a limitation, we can still observe in Figures 3.2 and 3.3 that there is a more complicated pattern with HMRC.
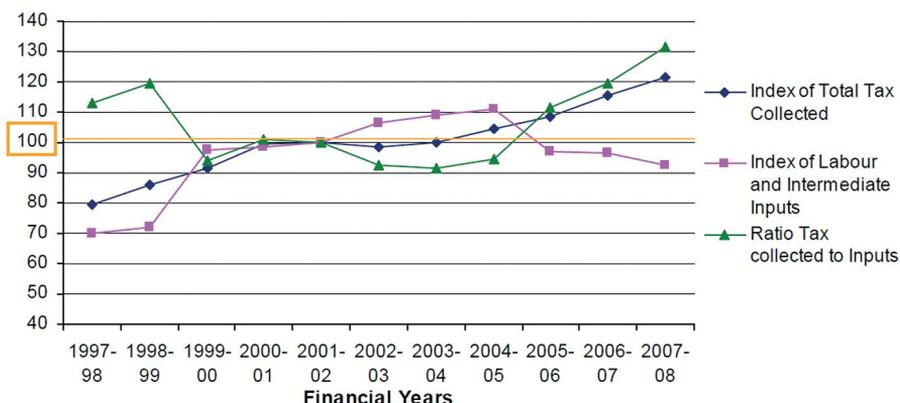
Volume (2001/02=100)

Figure 3.3. The ratio of the deflated amount of tax collected to labour and intermediate outputs for HMRC and predecessor departments, 1997 to 2008
Source: Patrick Dunleavy.

The volume of tax demand had been steadily increasing, but for a long time the green line indicating productivity hovered at just around 100, and had not quite reached 100 in 2001. There was quite a big increase in staff from 1997 to the middle of the Blair Government period (1997–2007), but subsequently a reorganisation of the tax agencies in Britain resulted in a merging of Inland Revenue and Her Majesty's Customs and Excise to create HMRC. At that point, we can observe the staff savings being accumulated and the inputs coming down. Staff numbers, which had totalled about 115,000, were reduced to about 70,000 at the time of the merger.

During this period, the tax agency became a lot more successful in getting people to use online forms, and made some other big technological changes. These took a long time to come properly into effect, but eventually led to a sharp increase in productivity. The increases caused by the merger and the new technology continued until about 2011.

By 2011, there had been overcutting. The tax agency failed to answer 44 per cent of phone calls made to it, which had a negative effect on public legitimacy. So there was a period of long stagnation (which became a big growth period), there was continued growth for a few years, and then a few problems emerged.

If we look at this in relation to tax raised per staff member, or staff input and contractor input (Figure 3.3), we see essentially the same pattern—the green line remained static near the productivity line for quite a long time, and then following the reorganisation and the IT developments, it began to significantly improve.

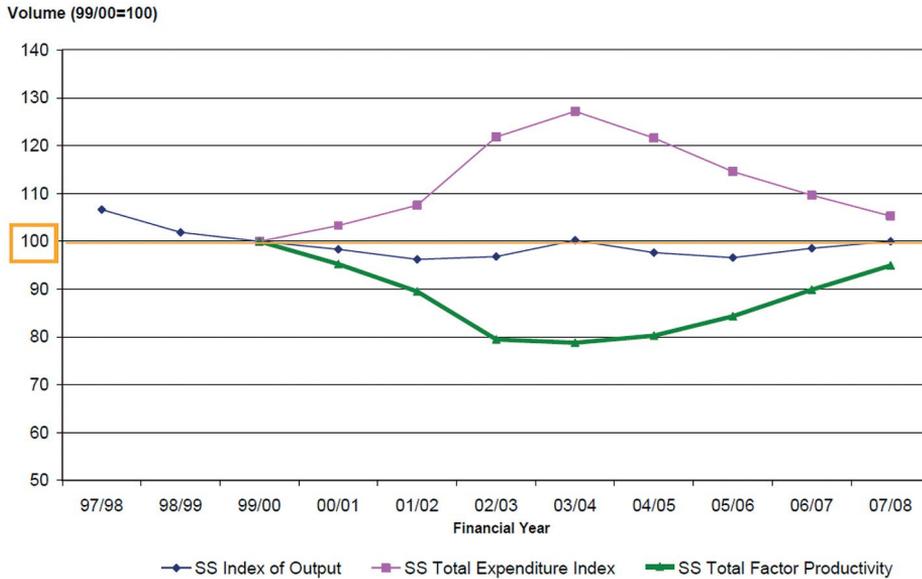# Department for Work and Pensions (WAP)

Volume (99/00=100)

Figure 3.4. Total factor productivity in UK social security, 1997 to 2008

Source: Patrick Dunleavy.

The third agency I will discuss is the Department for Work and Pensions, which is responsible for all the UK's social security. Figure 3.4 indicates that output throughout this period was fairly static, only going down a little as Britain came out of recession. However, there was a big increase in inputs (staff and contractors), and as a result the overall productivity slumped significantly in the period from 1999 to 2003–04.

This stemmed from two big decisions. First, there was a merger in 2001 (primarily a combination of the Department of Social Security and the Employment Service), after which the agency decided to fundamentally change the way in which it did technology. It moved away from paper, but not towards online services—the departmental leadership did not believe that welfare clients would ever use the internet, so they built the whole department around a telephone-based model. This tripled the number of phone calls it received. Unfortunately, however, most of these phone calls were not useful—they were just cluttering up the call centres. Clients were complaining about things, while staff did not know what was happening and could not help. The department made a very good effort to cut the number of phone calls over the next three years but, as can be seen in

Figure 3.5, it consequently used more postal services and more paper forms than it had in the beginning (although it did save some money by cutting face-to-face contacts).
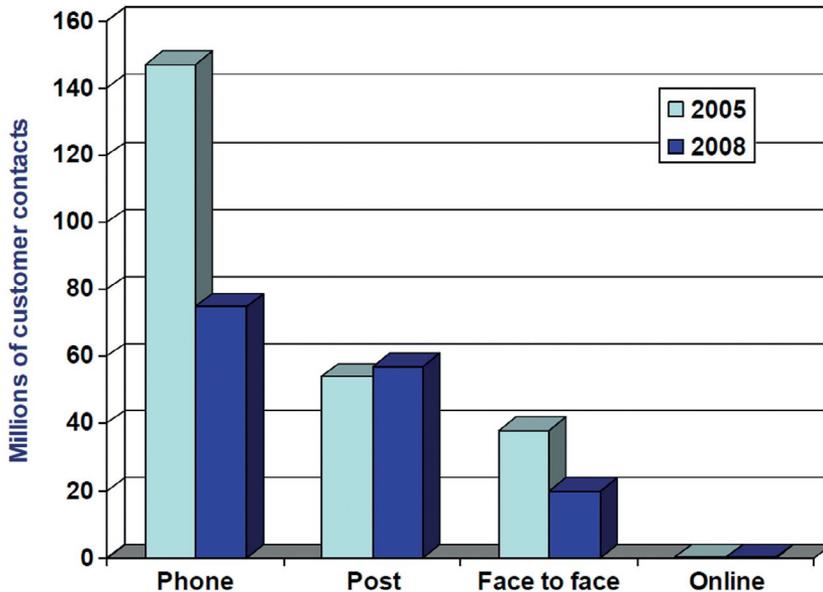

Figure 3.5. The changing pattern of the WAP's customer contacts, 2005 to 2008

Source: Patrick Dunleavy.

Note the online story shown in Figure 3.5; only 0.5 per cent of customer transactions were done online in 2005, and the same in 2008. By the time we conducted our study for the National Audit Office, 51 per cent of UK welfare recipients had broadband access, yet only 0.5 per cent were doing online transactions. Subsequently, the department has made a huge change in policy, moving towards what is called 'digital by default'—the vast majority of transactions to be conducted online. It was a belated move, but at least the department has now come into the twenty-first century. Fundamentally, the decision not to go digital was a strategic 'mis-decision': management miscalculated where society was going to go, made the wrong decision, and put a lot of effort into what was a doomed strategy.

Second, the Labour Government wanted to overhaul and modernise welfare benefits. They may have had very good reasons for that, and they certainly improved the flow of money to various disadvantaged groups, especially the elderly. However, the introduction of new policy requires more staff; and when the policy does not work terribly well, more people need to be hired. As a result,

policy change is associated with a large increase in staff inputs. The productivity measure does not tell us that it was not a good idea *socially*, but it does indicate that in simple productivity terms there was a big slump.

Factors influencing productivity

So what shapes productivity? What things work? What things don't work?

One thing that clearly works is going digital—taking it seriously, recognising that society is going digital and that 10 years from now we need to have essentially digital public services, born digital, delivered to citizens who are digital natives. It is not enough to layer a little bit of digital on top of paper-based or pre-existing twentieth-century systems. We must think about how we can do digital in a really radical way.

In the same vein, the second thing that works is thinking about restructuring services, really studying our services, really understanding them, really seizing control of them, really trying to change them and deliver public value as cheaply and as effectively as we can—including making major investments.

What *does not* work is bringing in consultants who do not adequately know the agency's service, and so give them off-the-cuff solutions that worked somewhere else but may not be appropriate for that agency. On its own, outsourcing does not do anything for productivity. An agency might try outsourcing for one of any number of reasons—because the minister tells them to, because they have faith in economic solutions, because they want more diversity—but we cannot expect it to improve productivity.

Going digital

It was very difficult and time-consuming to get enough good data on investment in IT—it took my co-author and me three or four years just to get the data and we have not got many data points. Figure 3.6 is not as good a graph as we would like, but it does provide some idea of how IT investment affects productivity. It shows a positive association between investing in IT now as a percentage of administrative costs, and what happens to productivity two years from now. The graph does not display very strong clustering but there is nonetheless evidence that it is working. Dunleavy and Carerra (2013) contains more detailed descriptions of this dynamic, including at the local level.
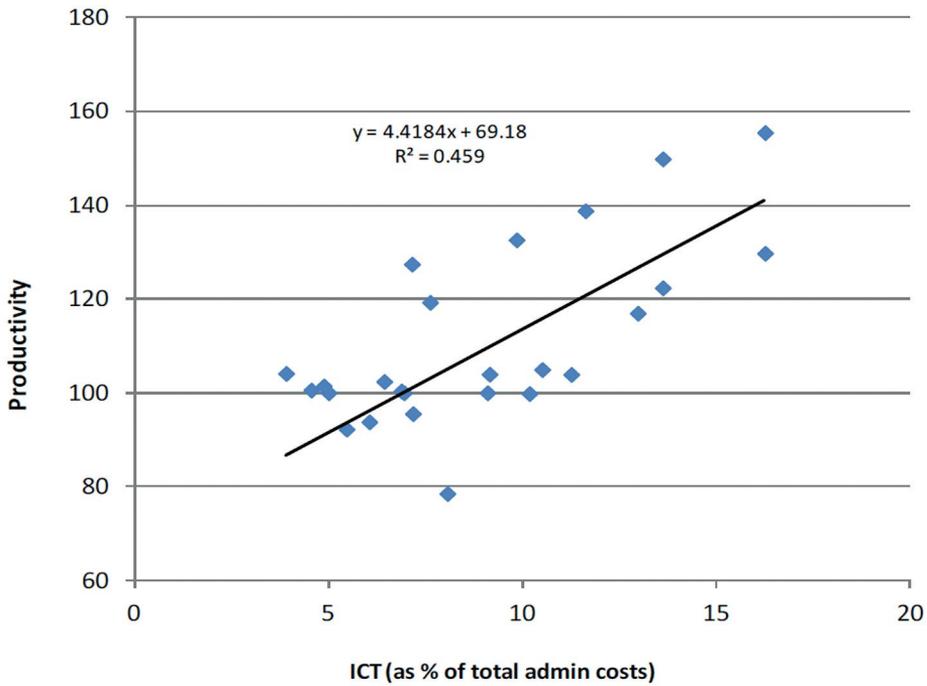
Figure 3.6. Productivity versus lagged ICT spending across UK welfare, tax and customs, 1999 to 2008

Source: Patrick Dunleavy.

Restructuring

The second thing that works is a major think-through and restructure of what an organisation is doing. The measure used in Figure 3.7 is investment in Private Finance Initiative (PFI) projects; when the Blair government came to power there was a lot of investment in new offices, new processes and new work processes. When we build or open a new building, we tend to reorganise the staff a lot more fundamentally than we would otherwise. Again, there is a positive relationship here—on the whole, the more new construction there is, the more we reorganise our work processes, the more our productivity tends to go up. There are a few troubling outliers here, but it is quite a strong clustering otherwise.
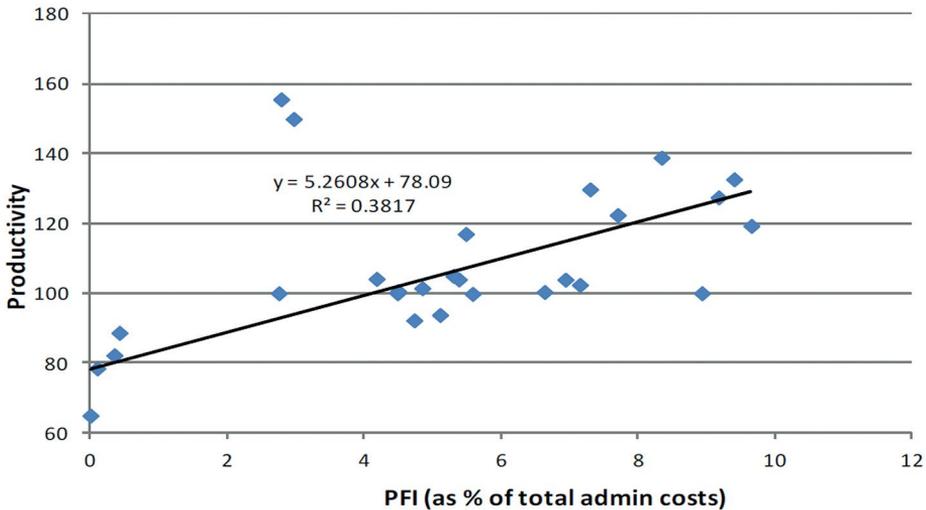
Figure 3.7. Productivity versus lagged construction PFI spending as proxy variable for major reorganisation; UK welfare, tax and customs, 1999 to 2008

Source: Patrick Dunleavy.

Consultants and outsourcing

We also looked at the impact of consultants, and it is very inconsistent (see Figure 3.8). There is no relationship between spending on consultants and productivity. If there is a relationship, it is negative.

So, why does outsourcing not work? It is because government service offices are highly imperfect and they are not going to stop being highly imperfect if two or three contractors are brought in. The markets created are oligopolistic. In Britain we have large problems with our IT sector—62 per cent of the market is dominated by the top contractor, and the top five contractors have 95 per cent of the market. There are usually only two or three tenders for any given contract, and the tenders are very expensive. The idea that more firms can bid is not feasible, because a firm needs to have a large governmental relations unit and a contracting unit just to understand the e-procurement system; this will always be the case. Contract specification works directly against productivity because an organisation needs to specify what it wants the contractor to do. It has to fix a whole service specification and then as needs change, and demand changes, and society changes, it has to go back to the contractor and renegotiate.
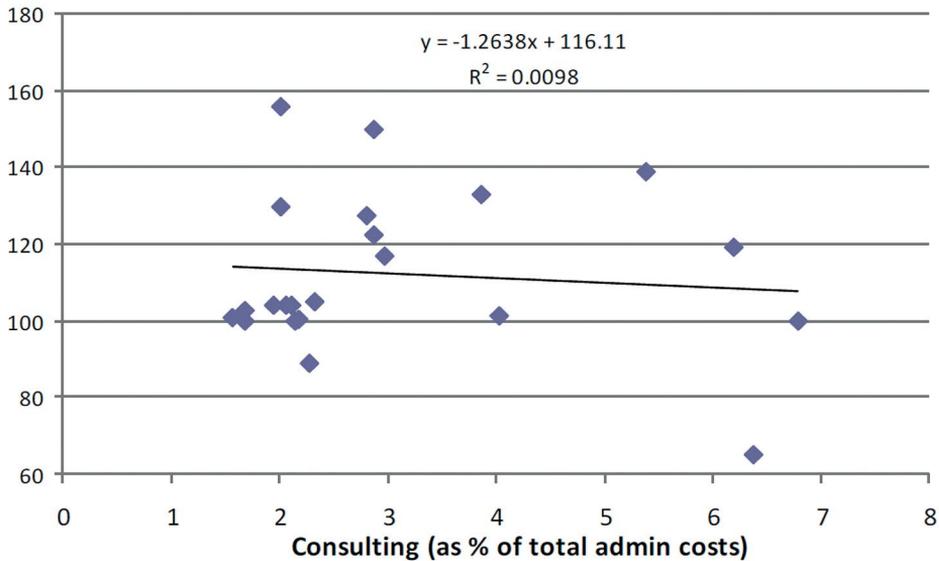
Figure 3.8. Productivity versus lagged consulting spending; UK welfare, tax and customs, 1999 to 2008
Source: Patrick Dunleavy.

In the British government IT market, the famous rule was 'six for one', meaning that if you had open competition for £1 million worth of business, you knew that over the life of the contract you would in reality get six times as much because of the renegotiation that would be forced by political or social changes. There are some outrageous examples of contract fixing. The current IT contract for British tax is an 18-year contract for IT services—and IT changes five times faster than normal life, so it is the equivalent of a century-long contract for a normal area of business.

Public servants also tend to use outsourcing in a very rational way—if we have better business to be attending to and there is something that we really hate doing, we tend to outsource it. This means that nothing changes in that area. The contractor will not want to change—as soon as we outsource it to them, they will want to freeze the technology and keep things exactly the same. This may seem irrational, because at the end of the contract they will have to re-tender, but it is actually cheaper for contractors to work that way.

When we have genuine cost reductions, it is likely to be a signal that lower contractor prices require a sacrifice in performance. For example, the UK progressively outsourced all its hospital cleaning. Fifteen years later, the UK had a hospital-acquired infection rate 15 times greater than the European average, and

40 times greater than field-leader Denmark. We then had to re-internalise this area and also reallocate work back to a few expensive, oligopolistic contractors, in order to fix the problem.

One final note—*contestability* is a great word, and it may do some good when trying to introduce product diversity, or when attempting to engage different kinds of contractors. The arrival of mutuals might make a difference, but keep in mind that mutuals only have 1/70th of the outsourcing market in the UK, so they are not a serious threat to the big outsourcers yet. On the whole, outsourcing contestability will not grow government productivity.

Eight steps to improving public sector productivity

So what can we do? From our recent book and the work we have been doing for the last 20 years, we can draw four immediate steps and four longer-term steps.

Short-term recommendations

1. Talk about it, look at it, measure it, and try to improve it. If we are not doing that with our organisation, we're not in the game. We need to have a viable minimal definition of productivity; we need to keep it simple. Perhaps as we get better we can make it more elaborate by bringing in more quality weights and starting to think about outcomes. But that's optional, and my recommendation is to keep it simple, try to get some measured productivity going in our organisation, and then build it up over time. Keep the definition stable, rather than constantly changing how we define everything. Try to obtain a time series and then build up a narrative about what it is that drives productivity in our organisation.

2. Public sector innovation is the lifeblood of productivity. Early private sector estimates said that 60 to 70 per cent of productivity gain in an industry was produced by displacement—that is, by customers leaving a firm that was less efficient or had an inferior product and moving to a firm that was more efficient and had a better product. That ratio has come down over time, and nowadays home-grown innovation is far more important in private sector productivity gain. But 50 per cent of all productivity gain still comes from people moving from the worst firms to better firms, while the bad firms go out of business.

This is an effect that cannot be reproduced in the public sector. People have been trying to produce supplier succession in the new public management period for the last 25 years. These efforts are feeble because there can be no supplier succession, and most public sector agencies are immortal, and will continue to be immortal—that is the good news for public servants. Consequently, the public sector will not see the same displacement as the private sector, certainly not in the short term under current arrangements. This means that home-grown innovation produced by managers—particularly medium managers, lower managers and staff themselves—is the lifeblood of change in the improvement of productivity.

3. Take digital seriously. I know everybody thinks they are taking digital seriously, but government websites across the world are vastly out of date, few are using blogs or Skype, and Twitter use is still rare. The IT world changes constantly; government has to be proactive and stick with what citizens and companies are doing. We should be thinking about moving to a future in which digital services are the core of what we are doing, and other things are add-ons. In the area of personal services, or professional services at the local or state levels, there must still be a lot of digital enhancement occurring.

4. Bring the workers with us. The people who know how to improve things are the people who are doing the job. But why do they not now tell us how to improve things? My feeling is that they do not trust us. There is a concern that if they offer up information, management will say, 'Thanks very much! Okay that's a great idea. We'll start doing that, and then we'll fire one of you, or we'll speed up the work process.'

This is the fundamental problem inside public sector organisations. I have been to organisations where consultants have found that staff are using two different computer systems—an old one and a new one. The new one is much more time-consuming than the old one, but management are really dependent on their expert workers, who are doing it on the old system in unrecognised ways. And this is the only way that things are being done. These are large organisations that have spent a lot of money on change, but they have spent it poorly because they have not talked to their workforce.

Management consultants will tell you their job is to go into organisations, talk to the managers who do not really know how the organisation is working, then talk to the workers that the managers have not bothered to consult (or feel that they cannot because it is 'beneath their dignity' to go and spend time with the workforce). The consultants then write down what the workers tell them, and tell it to the managers—that is what management consultancy is all about, in a nutshell.

Longer-term recommendations

1. The question is, can we have genuine demand transfers across suppliers? Can we get genuine supplier succession, genuine competition or contestability? I think we could if we had public sector suppliers who could scale up their services; who could move from one area to another and enlarge. More mixed public/private competition could also improve the situation, and mutuals may help in a small way here.

2. We must try to persuade the public to stop supporting big state policies from the late twentieth century that really are not working anymore, and to have enough trust in politicians and managers to embrace the new kinds of service patterns that would work better.

3. We need to exploit digital to focus public services, again, on free services. We can get free services from a company like Google—but can we get a genuinely free service from the state? There is a lot of potential to develop this area in the next few years.

4. Citizen co-production: we have tended to focus on producing a fully specified, very finished public sector product—one which is fully controlled. We need to move towards producing more part products that citizens can use for their own purposes.

But these are all longer-run changes. Nevertheless, there is enormous potential in them, especially in difficult areas such as health, social care and education.

References

Atkinson, Anthony B. 2005. 'Measurement of UK government output and productivity for the national accounts.' *Journal of the Statistical and Social Inquiry Society of Ireland* 34: 152–60.

Dunleavy, Patrick and Leandro Carrera. 2013. *Growing the Productivity of Government Services*. London: Edward Elgar.

Osborne, Stephen (ed.). 2002. *Public Management: Critical Perspectives*. London: Routledge.