



THE IMPACTS OF THE NEW PUBLIC
MANAGEMENT IN EUROPE: A META-
ANALYSIS

COCOPS Working Paper No. 3

Christopher Pollitt

Sorin Dan

22 September 2011



EUROPEAN COMMISSION
European Research Area



Funded under Socio-economic Sciences & Humanities

About COCOPS

The COCOPS project (Coordinating for Cohesion in the Public Sector of the Future) seeks to comparatively and quantitatively assess the impact of New Public Management-style reforms in European countries, drawing on a team of European public administration scholars from 11 universities in 10 countries. It will analyse the impact of reforms in public management and public services that address citizens' service needs and social cohesion in Europe. Evaluating the extent and consequences of NPM's alleged fragmenting tendencies and the resulting need for coordination is a key part of assessing these impacts. It is funded under the European Commission's 7th Framework Programme as a Small or Medium-Scale Focused Research Project (2011-2014).

About the Authors

Christopher Pollitt is BOF/ZAP Research Professor of Public Management at the Public Management Institute, Katholieke Universiteit Leuven.

Sorin Dan is Researcher at the Public Management Institute, Katholieke Universiteit Leuven.

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement No. 266887 (Project COCOPS), Socio-economic Sciences & Humanities

ISSN 2211-2006

THE IMPACTS OF THE NEW PUBLIC MANAGEMENT IN EUROPE: A META-ANALYSIS

Abstract

The purpose of this working paper is to rehearse the approach, methods and preliminary findings of the first stage of the EU Seventh Framework programme project *Co-ordinating for Cohesion in the Public Sector of the Future* (hereafter COCOPS – see www.cocops.eu). This first stage (Work Package 1) was tasked to produce *a meta-analysis of the impacts of New Public Management (NPM) reforms across the EU*. In this paper, therefore, we will first set out what we mean by (respectively) NPM, impacts and meta-analysis, and then go on to review the substance of our findings thus far. The database currently has over 500 documents, including academic articles and books, government reports and consultancy studies. It covers 23 EU member states, plus Norway, Switzerland and Croatia, plus the European Commission itself. Its contents are not yet absolutely finalised, but our collection, coding and quality checking have reached the point where we do not expect the broad conclusions to change very much. This is therefore an appropriate time to issue a Working Paper.

Keywords

New Public Management, public sector reform, impact, meta-analysis, European Union.

THE IMPACTS OF THE NEW PUBLIC MANAGEMENT IN EUROPE: A META-ANALYSIS

1. Introduction

The purpose of this working paper is to rehearse the approach, methods and preliminary findings of the first stage of the EU Seventh Framework programme project *Co-ordinating for Cohesion in the Public Sector of the Future* (hereafter COCOPS – see www.cocops.eu). This first stage (Work Package 1) was tasked to produce a *meta-analysis of the impacts of New Public Management (NPM) reforms across the EU*. In this paper, therefore, we will first set out what we mean by (respectively) NPM, impacts and meta-analysis, and then go on to review the substance of our findings thus far. This definitional exercise was anything but simple, and the interpretation of our results rests significantly on a number of significant definitional and operational choices that we made (and were, in fact, obliged to make, in one way or another). These choices are discussed in sections 3-7 below. We have distinguished between impacts on *processes*, *outputs* and *outcomes*. Within those broad categories we have paid particular attention to impacts on efficiency, effectiveness, and the attitudes of those who use public services. We have also looked for evidence of impacts on *social cohesion*, which for our particular purposes we have defined as having to do with equality of access to services and with the solidarity and commitment of public servants themselves.

The findings reported are preliminary in at least two senses. First, they are a first, quick examination of a not-yet-absolutely-complete database. We already have more than 500 documents uploaded, including academic studies and official reports, but we are still occasionally finding relevant new material, so some of the descriptive statistics in our final report will have to be updated. Second, there is only room here for an analysis of some of the rather larger (and perhaps) cruder dimensions of the database. In subsequent papers we intend to report more focused and detailed analyses of specific issues.

As a working paper the *raison d'être* of this text is to stimulate comment. This is work in progress. We will be grateful for questions and observations, whether they are made at a presentation of this paper or by email or other means.

2. Defining NPM

Defining NPM is the first step but in itself is not at all easy. In its origins it is strongly associated with UK Prime Minister Thatcher and US President Ronald Reagan, and with the New Zealand Labour government of 1984. Neither Mrs Thatcher nor Ronald Reagan were

any friends of the ‘planning’ approach which had been the orthodoxy in the US and UK public sectors of the 1960s and early ‘70s. During their periods in power in the 1980s they, and many of their advisers, favoured what they considered to be a more ‘business-like’ approach. Gradually, partly through doctrine and partly through trial and error, this general attitude crystallized into a more specific set of recipes for public sector reform. By the early 1990s a number of influential commentators appeared to believe that there was one clear direction - at least in the Anglophone world. This general direction was soon labelled as the New Public Management (NPM) or (in the US) Re-inventing Government (a seminal article here was Hood, 1991). A pair of American management consultants, who wrote a best-seller entitled *Reinventing government* and then became advisers to the US Vice President on a major reform programme, were convinced that the changes they saw were part of a global trend. They claimed that ‘entrepreneurial government’ (as they called it) was both worldwide and ‘inevitable’ (Osborne and Gaebler, 1992: 325-328). At about the same time the Financial Secretary of the UK Treasury (a junior minister) made a speech claiming that the UK was in the forefront of a global movement:

“All around the world governments are recognising the opportunity to improve the quality and effectiveness of the public sector. Privatisation, market testing and private finance are being used in almost every developing country. It’s not difficult to see why.” (Dorrell, 1993)

The increasingly influential Public Management Committee of the OECD came out with a series of publications that seemed to suggest that most of the developed world, at least, was travelling along roughly the same road. This direction involved developing performance management, introducing more competition to the public sector, offering quality and choice to citizens, and strengthening the strategic, as opposed to the operational role of the centre (see e.g. OECD, 1995). Whilst it is now fairly clear that the whole of the world was *not* following the same path (Pollitt and Bouckaert, 2011) it remains true that NPM ideas spread very widely, and are often still seen as the most obvious route to modernization.

There have been many definitional disputes and ambiguities about exactly what the key elements of this widespread trend were supposed to be: 'There is now a substantial branch industry in defining how NPM should be conceptualised and how NPM has changed' (Dunleavy et al., 2006b: 96; see also Gow and Dufour, 2000; Hood, 1990; Hood and Peters, 2004). For the purposes of this meta-analysis we will assume (like Dunleavy) that the NPM is a two-level phenomenon. At the higher level it is a general theory or doctrine that the public sector can be improved by the importation of business concepts, techniques and values. This

was very clearly seen, for example, when the then US Vice President personally endorsed a popular booklet entitled *Businesslike government: lessons learned from America's best companies* (Gore, 1997). Then, at the more mundane level, NPM is a bundle of specific concepts and practices, including:

- Greater emphasis on 'performance', especially through the measurement of outputs
- A preference for lean, flat, small, specialized (disaggregated) organizational forms over large, multi-functional forms (e.g., semi-autonomous executive agencies – see Pollitt et al, 2004)
- A widespread substitution of contracts for hierarchical relations as the principal coordinating device
- A widespread injection of market-type mechanisms (MTMs) including competitive tendering, public sector league tables, performance-related pay and various user-choice mechanisms
- An emphasis on treating service users as 'customers' and on the application of generic quality improvement techniques such as TQM (see Pollitt, 2003: chapter 2)

Dunleavy et al. have usefully summarized this as 'disaggregation + competition + incentivization' (Dunleavy et al., 2006a). Roberts (2010) has added a further interpretive twist by characterizing many NPM reforms as exhibiting 'the logic of discipline'. By this he means a way of thinking about the organization of government functions that has two major components:

- Scepticism about conventional democratic politics because it tends to produce shortsighted, unstable and self-interested policies
- Optimism that if certain subject or activities are legally removed from everyday politics (i.e. if elected politicians are 'disciplined' so that they cannot constantly intervene and interfere) then more stable and farsighted policies will be possible

Thus, for example, central banks should be given independence. Executive tasks should be hived off to contract-like agencies, local economic development should be removed from the responsibility of elected local councils and given to special-purpose businesslike bodies. Roberts says that this liberalization agenda revolved around ideas of depoliticization, autonomization and discipline (2010: 140). Experts and managers will (it is assumed) generate more stable, prudent and realistic policies and decisions than short-term, electorally hypersensitive politicians.

It would be quite mistaken to assume that this formula was necessarily internally consistent. As a number of commentators have noted, there is some tension between the different

intellectual streams that feed into the NPM, particularly between the economic, principal-and-agent way of thinking, which is essentially low trust, and the more managerial way of thinking which is more concerned with leadership and innovation – and more trusting of the inherent creativity of staff, if only they are properly led and motivated (Pollitt, 2003: 31-32). The former stream emphasizes the construction of rational systems of incentives and penalties to ‘make the managers manage’. The latter emphasizes the need to ‘let the managers manage’ by facilitating creative leadership, entrepreneurship and cultural change. Other writers have drawn a distinction between ‘hard’ and ‘soft’ versions of NPM (Ferlie and Geraghty, 2005). The hard version emphasizes control through measurement, rewards and punishment, while the soft prioritizes customer-orientation and quality, although nevertheless incorporating a shift of control away from service professionals and towards managers. This seems to map quite closely onto the low trust/high trust tensions mentioned above. As for the ‘logic of discipline’ Roberts comes to the conclusion that, in many cases, depoliticization turned out to be anything but politically neutral, organizational autonomy proved elusive (rather organizations swapped one framework of control for another) and ‘discipline’ was interpreted differently in different cultures (Roberts, 2010: 140-146).

3. Defining the domains of NPM reforms

For practical reasons of research capacity the COCOPS team decided at an early stage not to try to pursue every NPM reform in every EU country at every level. We limited our search in various ways:

- a) We excluded acts of outright privatization from the Work Package 1 meta-analysis (although privatization will receive some attention in some of the later work packages). We define privatization as the sale of publically-owned assets to the private sector. Therefore corporatization within the public sector remains within our scope, as does contracting out and the creation of autonomous state agencies
- b) We decided to concentrate on central government reforms. We have not attempted to cover reforms of sub-national tiers of government, although here and there, selectively, we include studies of local government if they offer a particularly clear account of the impacts of NPM reforms
- c) We dealt with social cohesion in a very narrow way. Basically, we focused only on two aspects. First we noted any evidence of how reforms might have influenced equity (especially equity of citizen access) in public services. Second, we also paid attention to any elements in the documentation which reported on changes in public service staff morale or attitudes towards their work. We see such changes as at least connected to the idea of cohesion, in the sense that, if morale plunges downwards it is reasonable to assume that the social cohesion of the public service is adversely

affected. Equally if, by contrast, reforms cause a marked increase in positive attitudes among staff, then that should be a factor in favour of organizational cohesion within the public service. We should note, however, that positive attitudes among staff are not necessarily or invariably correlated with organizational cohesion. So this is a very tentative exploration. Social cohesion is a complex and slippery concept (Chan et al., 2006) and will be addressed in a more extensive way by some of the later work packages of the COCOPS project

- d) We treated some EU member states more intensively than others. Table 1 (below) shows the level of scrutiny we gave to the different states. This was determined by a variety of factors, including the knowledge, linguistic abilities, and personal networks of the partners in the COCOPS team. Obviously it was easier to survey the literature of a country where a COCOPS partner was based, and obviously it was much more difficult in countries where none of our partners had command of the language (e.g. Greece). Our coverage, however, was equally shaped by some theory-related principles of selection. We aimed at - and achieved - coverage in depth of at least two countries in each ‘cultural group’. By cultural group we mean countries that shared important political histories and norms – elements that continue to shape their approach to public administration. The main groups we identified are all widely recognized as such in the academic literature, viz. the Nordic group, the ex-Napoleonic Mediterranean states, the CEE and the ‘big three’ (France, Germany and the UK - each of which has developed a strong style of its own, which has influenced other countries closely connected with it). Further discussion of these groupings can be found in, e.g., Lynn, 2006; Ongaro, 2009; Painter and Peters, 2010.

Table 1. Level of coverage of NPM reforms in COCOPS work package 1

State	Level of treatment in WP1
Austria (AT)	In depth
Belgium (BE)	In depth
Bulgaria (BG)	Not covered
Croatia (HR)	Medium
Cyprus (CY)	Not covered
Czech Republic (CZ)	Medium
Germany (DE)	In depth
Denmark (DK)	Medium
Estonia (ES)	In depth
Finland (FI)	In depth
France (FR)	In depth

Greece (GR)	Light
Hungary (HU)	In depth
Ireland (IE)	Light
Italy (IT)	In depth
Latvia (LV)	Light
Lithuania (LT)	Medium
Luxembourg (LU)	Light
Malta (MT)	Not covered
Netherlands (NL)	In depth
Norway (NO)	In depth
Poland (PL)	Light
Portugal (PT)	Medium
Romania (RO)	In depth
Slovakia (SK)	Medium
Slovenia (SI)	Not covered
Spain (ES)	In depth
Sweden (SE)	Medium
Switzerland (CH)	Medium
United Kingdom (UK)	In depth
European Commission (EC)	Medium

We should also note that defining what is and what is not a public organization is by no means always straightforward (see for example, the complexities encountered in CEEP, 2010 or Clifton and Díaz-Fuentes, 2010: section 2). By focusing principally on central government bodies we hope we have, however, avoided the worst of this.

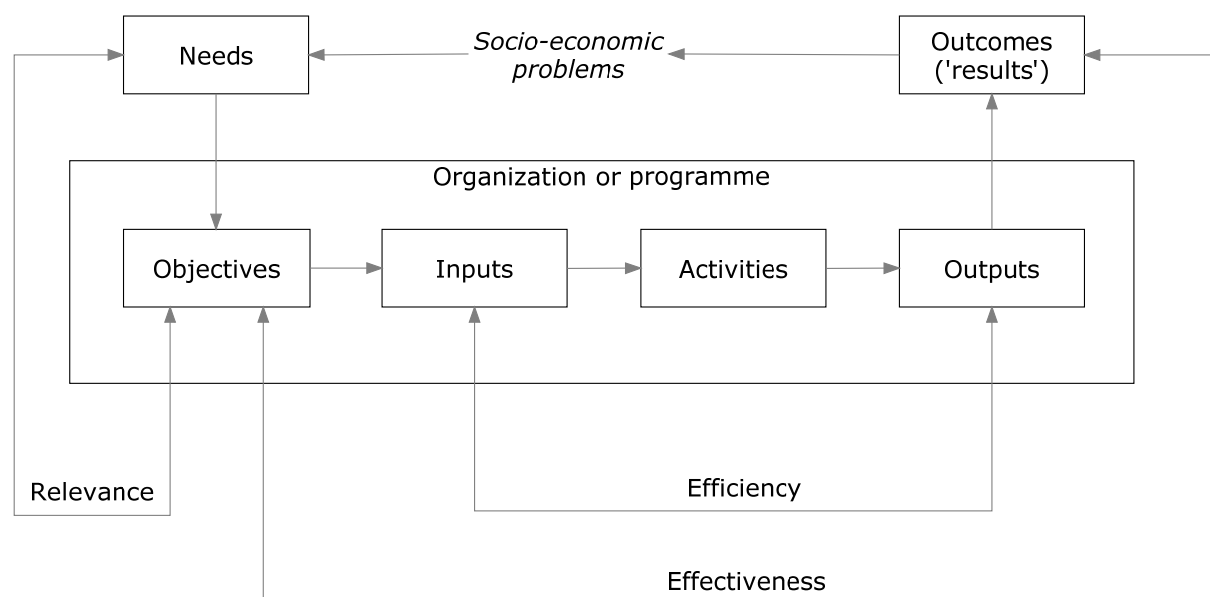
Finally, it is important explicitly to state that NPM reform and public management reform are by no means the same thing. It is necessary to make this pronouncement because there has occasionally been a tendency, perhaps especially among continental European commentators, to see anything that is labelled ‘management reform’ in the public sector as being some species or aspect of NPM. We find this loose and confusing – to paraphrase Aaron Wildavsky, if NPM is everything, then it is nothing. Our position is quite clear: a) NPM reforms partake of several or all of the characteristics identified above, b) there are many other types of reform going on which cannot be classified as NPM (for example, ‘joined-up government’ or collaborative networking) and c) very important though these other types of reform undoubtedly are, this paper is not concerned with them. Our focus is on the impacts of NPM.

4. Assessing impacts

The COCOPS research proposal committed Work Package 1 to a ‘meta analysis of the impact of NPM on efficiency, effectiveness, quality and social cohesion’. Almost every term in this short phrase has led us into extensive internal debate, not least the notion of ‘impacts’.

‘Impacts’ are things that can reasonably be supposed to have been *caused* by the reforms. They could be things very close to the reforms (for example, staff lost their jobs because some activities were contracted out) or much ‘further out’ (young people got better jobs because the quality of their education went up, because, in turn, the performance of the schools they attended was closely measured and publicized, so that teachers were obliged to improve the standard of instruction). In thinking about impacts we adopted a fairly standard model of policymaking (see Figure 1, below)

Figure 1. Performance: A conceptual framework



(Adapted from Pollitt and Bouckaert, 2004:13)

Although the conceptualization and the terms in Figure 1 are, as we just said, fairly standard, our repeated experience is that they are nevertheless very variably interpreted. In some cases normatively positive terms like ‘efficient’ or ‘effective’ are stretched in order to make an argument sound better. For example, government reports may claim an increase in ‘effectiveness’ when what they actually mean is that organizational staff had adopted better working methods - which in terms of Figure 1 would be just process changes, and nothing to do with outcomes (for an example of this see PA Consultants, 2002). Politicians, in particular, splash words like ‘efficient’ and ‘effective’ around like fresh paint, but some academics are also fairly loose in the way they use these terms.

We therefore need to say a few words about the relationships conceptualized and represented in Figure 1, and how we intend to use these terms. First, an organization or programme is conceived as a set of *activities or processes*. These would include organizational arrangements like the division of responsibility, the allocation of authority, the standard operating procedures, and so on. These procedures enable the organization to deploy and redeploy its resources (staff, money, buildings etc) which are collectively termed *inputs*. From these activities and processes the organization or programme then produces a set of *outputs*, which could, for example be lessons (in a school), licenses (from a licensing agency), medical treatments (from a hospital) and so on. These outputs are, in a sense, what the organization ‘gives’ to the outside world – to citizens, to civil society associations and to business firms. They are like messages, passing across the membrane that separates (on the one hand) the state from (on the other) the market sector and civil society. Outputs are invariably intended to produce desirable *outcomes*, beyond the organization or programme - so school lessons are supposed to produce educated students and hospital-provided medical treatments are supposed to produce the cure or alleviation of ill-health. *An outcome is something that happens in the world outside the organization and the programme: it is an effect ‘out there in the real world’*. Some analysts make distinctions between ‘intermediate outcomes’ and ‘final outcomes’, but we decided that such a distinction, though useful in some contexts, would be of limited value in our field of study, where outcomes measures of any kind are rather rare.

One final point to add about effectiveness is that one measure that is potentially very useful to policymakers is that of *cost effectiveness*. It is not shown in figure 1 because it is essentially the same relationship as shown for basic effectiveness, i.e. the degree of match between policy goals and outcomes. However, cost effectiveness adds a unit cost calculation, i.e. how much of the goal is gained for each unit of expenditure. Thus, for example, a measure that showed how many cancer deaths were avoided per £1m spent on a cancer screening programme would be a cost effectiveness measure (‘an additional £10M would save an additional 14 lives’, for example). Unfortunately, there are very few cost effectiveness measures in the field of public management reform.

The performance of organizations and programmes (the value of their activities) is usually thought of in terms of certain relationships between these inputs, outputs and outcomes. Thus the ratio between inputs and outputs is a measure of efficiency (or ‘technical’ efficiency or ‘X-efficiency’, or ‘productivity’). If you can get more outputs for the same inputs, you have achieved an efficiency gain. If you can maintain steady outputs while reducing inputs you have also achieved an efficiency gain - in both cases the ratio between inputs and outputs

improves. Effectiveness, however, is a different concept, which is usually conceived as the degree to which the outcomes match the original goals or objectives set for the organization of programme. So, if a hospital is created with the goal of reducing premature deaths from heart disease, and it performs operations and delivers treatments which do indeed reduce the incidence of premature deaths within its catchment population, then it has been effective.

As many writers have remarked, if goals are multiple, conflicting or ambiguous then it will be difficult to determine effectiveness which will, in effect, become a 'contested concept'. Unfortunately for the analysts, policy goals frequently *are* multiple, conflicting or ambiguous, not least because that is what politicians may need to get sufficient agreement to launch the policy in the first place. [This certainly includes public management reforms which are often claimed to be all things to all men – to save money, raise service quality, increase effectiveness, etc, etc.]

Although this is a helpful and widely-used conceptual framework it also raises a number of questions. One is - where does public opinion fit in to all this? Our answer is that we *have chosen to treat changes in public opinion (e.g. changing levels of satisfaction or trust) in respect of public services as an outcome*. Clearly it is not an output, because it is not 'made' by the organization or programme - rather it is a reaction to or assessment of that output (among other things). However, it is a different kind of outcome from some other outcomes. Whereas there is something concrete and clearly countable about, say, reductions in the number of road deaths or increases in the percentages of students passing a national test, satisfaction and trust appear to be more subjective - and volatile. That does not make opinions any less important but it does imply that analysts will need to treat such data with considerable interpretive caution. Note that citizen reactions to public services, although *in themselves* a kind of outcome, may refer to any stage in the policy process. In other words, the public may have strong feelings about cuts in health service budgets (which are inputs) or about improved queuing systems in the benefits payment office (which is a process change) or about new kinds of school lessons (which are outputs) or about the way in which the environmental health service has allowed a local river to become visibly polluted (which is an outcome).

Note also that effectiveness and efficiency by no means always move together. It is perfectly possible to carry through reforms which improve effectiveness but which do not affect efficiency, or which even reduce efficiency. Similarly, it is wholly conceivable that one can make changes that will increase efficiency, but also lead to some loss of effectiveness. For example, if a university hires top professors instead of junior lecturers to teach its courses, the

effectiveness of its teaching may go up, but, *ceteris paribus*, with a higher salary bill and only the same number of lessons, its efficiency will go down.

Another point to be born in mind is that, increasingly, official reports and evaluations, as well as academic studies, make use of complex indices of performance, which combine two or more elements (see e.g. Audit Commission, 2009; Pollitt, 2011; Van de Walle et al., 2008). These aggregated indices can be very useful to busy decision makers or to non-specialists and citizens, but they may also (deliberately or inadvertently) conceal underlying trades-off between two or more desirable values (e.g. equity versus efficiency). They can also give spurious precision to judgments which are more correctly seen as hedged about by quite wide brackets of uncertainty (Jacobs and Goddard, 2007). Furthermore, from our point of view they pose classification problems. In Audit Commission 2009, for example, local authority performance is classified as ‘excellent’, ‘good’, ‘poor’ (etc.) where each of these categories is in practice a complex aggregation of measures of output, assessments of strategic planning and capacity in the authorities concerned, and a number of other elements. How then do we classify a shift from, say, ‘good’ to ‘excellent’? It could be due to an increase of effectiveness, or efficiency or strategic planning capacity, or any combination of these things. We have no short solution to this, and have merely attempted to pick out from such aggregates the separate elements – where we can.

The above approach to impacts can operate on very different levels. One study may deal with the impacts of a particular organization, another with the impacts of a reform programme which embraces several or many organizations, and a few studies actually try to capture changes right across the civil service or even the entire public sector. When we come to analysis of the contents of the database we will usually need to distinguish between these different kinds of study.

Apart from process changes, efficiency and effectiveness, our terms of reference required us to pay attention to two dimensions of performance not mentioned on figure 1 – ‘quality’ and ‘social cohesion’. Both terms appear quite frequently in policy as well as academic debates, but the sad truth is that both terms are even more subject to multiple and vague definitions than ‘efficiency’ and ‘effectiveness’. ‘Quality’ is a notorious term, in that its positive normative connotations mean that it has been used to cover almost any aspect of public services. It may be used to cover the views of service producers and experts, or the opinions of service users/consumers. The well-known reform approach known as Total Quality Management (TQM) rests heavily on the opinions of consumers to define quality (although some of these consumers may be ‘internal customers’ rather than the citizen in the street).

‘Quality’ has certainly been used to refer to all manner of combinations of processes, outputs and outcomes - which is one reason why it can be a source of confusion (Pollitt and Bouckaert, 1995: chapter 1). We therefore treat it as a multi-dimensional concept, the actual dimensions of which need to be defined in each particular context. Thus (for example) a study showing that a Total Quality Management (TQM) system had been installed at government organization X would count as a process change. [Installation alone does not guarantee a change in outputs, still less outcomes.] A deeper study, showing that service users at X had noticed faster more accurate service since the TQM system had been installed would count as an output. And a further study showing that the faster, more accurate service from X had led to healthier patients or catching more criminals (or whatever) would count as an outcome.

As for social cohesion, we have here another slippery, multi-meaning concept and one which, like ‘quality’ carries a distinct normative charge (Chan et al., 2006). ‘Cohesion’, like ‘quality’, presents as a Good Thing. It is a term much used within EU policymaking, but not nearly as prevalent in the Anglo-Saxon world. It usually has to do with avoiding or ameliorating gross socio-economic inequalities, on the grounds that such imbalances are both ethically and socially corrosive. Such inequalities can occur in a vast variety of forms (for only some of these, see Kearns and Forrest, 2001). They may be geographical (some regions or localities are left behind in development) or economic (an increase in income or wealth inequalities) or social (some parents are able to buy a high quality education for their children while others have no choice but to send theirs to substandard slum schools). Given our focus on NPM reforms, we decided to concentrate primarily on issues of access to public services. Did the reforms we are interested in generate increased inequalities in citizen access, or reduced inequality, or did they leave things more or less the same? It has to be said, however, that whilst there are a few important studies which examine this aspect (e.g. Boyne et al., 2003), most NPM literature is far more concerned with efficiency than with equity. Nevertheless, we look at what there is. A second aspect of cohesion that we pay some attention to is that of the attitudes and morale of public service staff. As we said above, we see falling morale as a possible indicator of some loss of cohesion within public sector organizations, and rising morale as a sign that cohesion may be increasing. Some of the reports we study do contain evidence of changes in morale (or related concepts) so we will include that in our analyses, although we are conscious that these connections (morale-cohesion) are by no means certain or automatic. Finally, with respect to figure 1, it should be noted that (again, like quality) threats to social cohesion can occur at any stage in the policy process – inputs, processes, outputs and/or outcomes. The *effects* of lowering cohesion - the ethical problems and the social corrosion - will, however, be regarded, whether intended or unintended, as a particular class of outcomes.

Figure 1 does not mention ‘impacts’ as such, so there is still a need to choose how widely to construe the notion. One could take a very tight, puritanical line and say ‘only outcomes should be regarded as impacts’. Or one could be more liberal and also allow outputs. Or one could adopt an even broader definition and include reform-induced changes in activities (processes) as well. To stretch even further back up the chain, and include inputs, seemed to us a step too far (if one puts more money into contracting out and less into in-house services, is that an ‘impact’ of the contracting out reform?). In the end we adopted the following guideline: *we will consider changes in outcomes, outputs and activities/processes as impacts, but will devote more analytic attention to outputs and outcomes than to processes*. The way we collected the data allowed us to distinguish studies in each of these three categories (i.e. studies which focus on activities/processes, output studies and outcome studies). Some studies obviously include data on more than one of these categories.

This is not by any means the end of the analytic problem. There is also the vexed issue of *attribution*. To be the impact of an NPM reform the outcome, output, or change in process, must appear to be the *result* of that reform, not of other developments which may be happening at the same time. There must be a plausible causality. This condition is, however, often hard to satisfy (in fact in the literature we have examined it frequently is not satisfied at all). A classic example would be shifts in public satisfaction, or trust, in government or in the civil service. If there is a shift (either upwards or downwards) it *could* be the result of a preceding public service reform. However, it *could* also be the result of a general well-being factor, linked to an economic boom, or to ephemeral shifts in party politics (a new leader, a scandal) or to pre-existing personal expectations or hopes, or any combination of these. Tracing shifts in public opinion to specific reforms can be very hard to do (Clifton and Díaz-Fuentes, 2010; James, 2010).

The issue of causality is, of course, both fundamental and controversial. Some species of contemporary social sciences will hardly admit the notion of causality at all. Others insist that causality is central to social science explanation, but adopt quite a restricted notion of what will count as causality (Pollitt and Bouckaert, 2009: chapter 9). In this work we broadly follow Kurki (2008: 210-234) in accepting a multi-dimensional concept of causality, in which material causes, formal causes, efficient causes and final causes can all play interacting roles in an explanation.

All in all, therefore, it can readily be appreciated that assessing the impacts of NPM reforms is very far from straightforward. It involves several layers of categorization and interpretation,

each of which has some effect on what is eventually ‘found’. Given our terms of reference, this is unavoidable. Terms like ‘quality’, ‘social cohesion’, and even ‘efficiency’ and ‘effectiveness’ are not stable, technical entities. They have entered common - or at least bureaucratic and political - speech, and have there been deployed in a fantastic variety of ways. Since COCOPS is very much engaged with the practitioner as well as the academic literature, we are obliged to deal with all this material, however messy or inconsistent it may sometimes be.

More generally, it is important to understand that the body of work contained within the database derives from a mixed, increasingly multinational community, consisting of academics, public servants, management consultants and politicians. It is therefore unsurprising that the reasons for becoming engaged with NPM reforms have differed. Some participants want to find the best way forward – reforms that will work to solve some real (or imagined) problems. Some want to justify a recent choice of a new direction – to defend a new policy against attacks from the political opposition or criticism from the media. Some wish to package and sell sets of ideas (‘best practice’, ‘the reinvention model’, etc.). Management consultants, ‘experts’ and governments all do this (e.g. Federal Ministry of the Interior, 2006; Kaufmann et al., 2007; Osborne and Gaebler, 1992; Prime Minister and Minister for the Cabinet Office, 1999). Some hope to sound progressive and look good at an election. And some - mainly the academics - simply want accurately to describe and explain what is happening or has happened in the world of reform.

This mixture of motives means that the dividing line between descriptive and analytical (‘is’) statements and normative (‘ought’) statements is frequently hard to find. The desire to understand and explain is often tangled up with the desire to promote and support a particular kind of reform. Those reading the literature therefore need to be especially sensitive to the likely interests of the author(s), to unspoken assumptions, to the strength of evidence in relation to the size of the claims being made, and so on. This is what used to be called ‘source criticism’ and it is a vital technique for those who wish to investigate the literature on NPM. For example, a someone who researched NPM solely by visiting government websites would be likely to come up with a picture of what was going on that was both over-simple and over-optimistic. Even texts produced by academics cannot be assumed to be ‘neutral’, partly because many public management academics also work in consultancy and advice roles but also because the academic world is itself divided between competing theoretical and methodological camps (Pollitt and Bouckaert, 2009: chapter 9).

5. Actual selection of studies for uploading to the database

Studies were identified both by our own literature review and by calling on the expertise of our COCOPS partner institutions, and of a number of collaborators who kindly volunteered to search out material from their own countries. We were very grateful for this assistance, and we list these collaborators in Appendix A.

As for our own literature review, we began by checking everything that had been published since 1980 in what we considered to be the seven of the leading mainstream Anglophone public administration journals (they are all listed in the Thompson Index/SSCI), plus a few others:

- Journal of Public Administration Research and Theory (JPART)
- Public Administration
- Governance
- Public Administration Review
- International Review of Administrative Sciences
- International Public Management Journal
- Public Management Review
- Public Policy and Administration
- Evaluation
- Transylvanian Review of Administrative Sciences

In addition to these academic sources we trawled a number of promising organizational publication lists, including those of:

- OECD
- SIGMA
- World Bank
- National government websites on central government reforms (in those countries where they are available)
- National audit office performance audits (in those countries where they are available)
- Civil society organizations in Central and Eastern Europe

Alongside this literature search, as mentioned above, we worked either directly or through our COCOPS partners with a number of collaborators. From these sources we gained citations of official publications in the countries concerned, as well as academic publications in non-Anglophone journals. The COCOPS partners were in Estonia, France, Germany, Hungary, Italy, the Netherlands, Norway, Spain and the UK. The volunteer

collaborators were mostly in CEE and the Baltic states but also in Italy, Portugal and Switzerland. This opportunity provided us with wider access to studies and better insight into the specifics of the evaluation of impacts of NPM reforms in the respective states, but it also challenged our attempts to standardize the coding and analysis process. It also showed us even more clearly how slippery the concepts that we operated with were and how challenging the goal of assessing impacts of NPM reforms was.

We developed a set of guidelines and definitions of key terms for our collaborators, who then provided the documents and a summary of key findings as they pertained to impacts. In a second stage, either we or our COCOPS partners checked the documents and data for accuracy and relevance and uploaded into the SharePoint database those that met the criteria for impact assessments of NPM-type reforms. In rare cases, some of the collaborators uploaded the documents directly into the database as it was the case for Portugal.

The slipperiness of the interpretive problems necessarily involved in studying NPM impacts across the Europe meant that, despite extensive discussions of the foregoing issues across all the partners in the COCOPS research team, and the development of explicit guidelines for study selection, when it came to actually uploading selected documents into the database, some inconsistencies appeared. For example, one partner uploaded a paper entitled ‘E-government and the transformation of public administrations in EU countries: beyond NPM or just a second wave of reforms?’ (Torres et al., 2005). The title certainly sounds relevant, but closer inspection revealed that a) the paper dealt solely with sub-national governments and b) that its relation to NPM as we have defined it was vestigial (it took a somewhat ambiguous position, arguing that both that e-government was a kind of successor to NPM, but also that it was a potentially transformatory route to something regarded as quite different, namely ‘governance’). Another example would be studies that discussed the logical coherence or underlying philosophy and values of a particular set of NPM reforms. On the one hand it could be said that such studies contain no empirical data on concrete impacts, and should thus be excluded. On the other it could be argued that this type of study can identify certain likely or logically necessary consequences of NPM, and therefore deserved to be included. Some studies of this type were in fact submitted to the database, but we have excluded them from most of our analyses, on the grounds that the discussion of ‘impacts’ really does require some empirical data and cannot just rest on *a priori* conceptual analysis.

How did we handle these inconsistencies? Basically, at the first stage we allowed all papers put forward by our COCOPS partners to be entered into the SharePoint database. However, subsequently we ourselves reviewed most of the studies and were able to exclude some from certain sections of the meta-analysis (see next section). We also checked the coding of the metadata for inconsistencies, and eliminated these as far as we were able. We were therefore fairly liberal in admitting studies to the database (out of a desire not to miss anything important) but rather strict in our subsequent interpretation of what these studies actually showed.

6. Meta-analysis

COCOPS is committed to produce a meta-analysis, but what is a meta-analysis? Again, there is no single view of this, and definitions vary considerably.

Probably the dominant conceptualization is that prevalent in healthcare policy and medicine. Here a meta-analysis is a very strict set of statistical procedures by which the results of many different primary studies (typically randomized clinical trials) are compared so as to yield an overall average estimate of the net effect of the drug/procedure/treatment (Cooper, 1998). This is absolutely *not* the type of meta-analysis we have attempted – and with the kind of data that is available in the public management literature it would in any case be impossible.

An intermediate type would be to still attempt a statistical averaging of results, but to admit studies that were not experimental, such as an ordinary survey of managers' opinions, or of citizens' perceptions of changes in service quality. This is feasible and, indeed, has been attempted (Andrews et al, 2011). This approach to meta-analysis yields interesting results but it necessarily restricts itself to a certain type of study – a type of study which, until now at least, has represented only a small fraction of the many and various writings about NPM and its effects. Thus one achieves a relatively precise answer, but on the basis of excluding the majority of both academic and practitioner studies. This did not recommend itself to us, at least partly because studies of the requisite type have been largely confined to the USA and the UK (see Andrews et al, 2011, Table A.1) and therefore we would not only have been drawing on just one small part of the literature, we would also have been excluding many EU member states entirely.

Our version is therefore more relaxed still. It includes some studies of the type that Andrews et al reviewed, but many other types besides. It aspires to be what Pawson has termed 'realist synthesis', although at times it may be closer to 'narrative review' (Pawson, 2002; Pawson et

al., 2005). Although different in many fundamental ways, classic meta-analysis, realist synthesis and narrative review all share a basic common aim:

‘Meta analysis is used to summarize and compare the results of studies produced by other researchers. A meta analysis is often done to establish the state of research findings in a subject: in this way it provides the researcher with an overview of what others are saying about the subject rather than another discussion of one or a few parts of the question, problem or issue’ (McNabb, 2010: 79)

One rule of classic meta-analysis which we have broken is the prescription that says ‘only include studies which are based on original empirical research’. Whilst most of the studies in our database are based (at least in part) on original empirical research we have also allowed in some which are more in the nature of synthetic, analytic overviews, i.e. analytic summaries of the work of others. The main reason is that these have sometimes contained useful interpretive insights which were not necessarily present in the original constituent studies that make up the synthesis. A lesser reason is that, in the public administration literature, both academic and even more practitioner, the line between original empirical research and secondary analysis is often far from clear. Many useful documents comprise an element of original research combined integrated with an analytic overview of other work. We saw no need to exclude this type of study from the database. Once again, we were fairly liberal in what we allowed in, but stricter in how we interpreted them.

In classic meta-analysis the intervention or programme is itself assumed to have causal powers. The hope is to identify the most powerful type of programme or intervention (Pawson, 2002: 341). But in the realist synthesis which we favour:

‘It is not programmes that work: rather it is the underlying reasons or resources that they offer subjects that generate change. Whether the choices or capacities on offer in an initiative are acted upon depends on the nature of their subjects and the circumstances of the initiative. The vital ingredients of programme *ontology* are thus its “generative mechanisms” and its “contiguous context”. Data extraction in a realist synthesis thus takes the form of an interrogation of the baseline inquiries for information on “what works for whom in what circumstances” ’ (Pawson, 2002: 342)

To put it simply, it is the *combination* of the *intervention* (policy or technique) and *context* which gives rise to specific impacts or results. So a given technique may produce quite different outputs and outcomes in different circumstances. For example, competitively contracting out refuse collection may, in a given context, turn out to be very successful, but that does not at all mean that the results of competitively contracting out brain surgery will be similarly positive. Contexts consist of complex assemblies of different elements, including interpersonal relations, organizational cultures, structures and procedures, legal frameworks and the political climate (Pawson et al., 2005). Because of our special interest in contexts we included a specific question in the meta-data section of SharePoint seeking to establish whether each study had identified specific contextual features and, if so, what they were.

7. Consequences of our definitions for the structure and scope of the study

Some of our decisions increased the size of our databank, and some reduced it. Our relatively generous definition of what a meta-analysis could be increased our ‘stock’. If we had adopted a strict medical style definition, in which we had only included randomized clinical trials (RCTs) of tightly defined reform-interventions, our database would have disappeared altogether! There is virtually no material of this kind with respect to major NPM reforms, in any EU country. Yes, there are some good outcome studies (although not that many, as we shall see later) but these are hardly ever based on the RCT model which is common in medical research. On the other hand our definition of impacts excluded a great deal of material, both academic and practitioner, in which NPM is discussed at length, and various claims about it are made, but no real evidence of specific effects on processes, outputs or outcomes is presented. Thus, for example, there is a fair amount of both French and English material which discusses (sometimes with considerable sophistication and elegance) the philosophical and ethical implications of NPM thinking, but which has no evidence of actual impacts. Similarly, there are spirited defences and attacks on NPM and ‘managerialism’ in the Anglo-American literature which score telling analytical points by relating it to broader neo-liberal doctrines or to a limited form of ‘technical rationality’, yet these accounts say nothing about anything concrete, and they do not advance any empirical evidence concerning impacts. We are not, of course, saying that these kinds of analysis are without value: we are simply saying that our focus is on impacts, and so they fall outside our terms of reference.

The results of our decision to include studies of processes (activities) are certainly debatable. This inclusion has considerably enlarged the range of material we are working with. And there is no doubt that some of the impacts - or claimed impacts - on processes are both interesting and important. For example, a number of studies claim that the creation, in the UK, of more than 130 executive agencies under the ‘Next Steps’ programme from 1988

resulted in a loss of capacity for policy co-ordination by ministries – that is a process impact (e.g. Office of Public Service Reform, 2002). On the other hand it has to be said that this finding, interesting though it undoubtedly is to ministers, civil servants and organization theorists, is still quite a long way from what the average citizen or media commentator might regard as a ‘result’. A result would be an outcome *resulting* from the decreased policy co-ordination – for example, if people had not been able to get needed hospital care and had therefore died because hospital beds were being needlessly occupied by recovering patients who could not be discharged because co-ordination between health organizations and social care organizations had broken down. That would certainly be an impact, classifiable as an outcome.

8. Some preliminary remarks concerning the overall characteristics of the materials

As we worked through the hundreds of documents in our database certain broad dimensions became apparent. Since these in an important sense ‘frame’ or condition the quantitative analysis provided in the next section it seems worthwhile setting them out here.

First, the proportion of documents that actually provide primary data concerning the *outcomes* of NPM-type reforms is very small. Such articles (e.g. Propper et al., 2008; Kelman and Friedman, 2009) are very valuable but also rare. It is probably no accident that these gems tend to occur most often in the historically two most professionalized and scientifically-based public services – healthcare and education. They also tend to come from those countries that maintain large performance databases, especially the UK and (beyond this study) the US. There are many more documents that look at outputs, but even these are not in a majority. At the other extreme, the number of secondary studies which discuss NPM reforms in a general way, citing other literature, but with no original data on outputs or outcomes is very large. [There are several reasons why this should be so, as indicated in section 4, above.] So, it might be said, the sheer size of the NPM literature is potentially misleading - only a small part of it directly pertains to our key question: what impacts has NPM had on outputs and outcomes? Or, to put the same observation in a more critical way, lots of people have had lots of ideas about NPM (many of them very interesting and persuasive) but rather few have gone out into the field to collect primary evidence concerning specific impacts on service users and the citizenry more generally.

Second, it is clear that the difficulties we have had in deciding on, and then keeping to, a consistent definition of activities/processes, outputs and outcomes are difficulties which are far from being ours alone. In both official reports and the academic literature the line between processes and outputs (in particular) has often been treated in a very plastic way. Thus, for

example, producing a report by a target date is often treated as an ‘output’, although we would normally see that as part of the internal process of government (few citizens are going to read or even hear of this report). Or again, achieving particular reorganizations are frequently treated as outputs or even outcomes - ‘we set up two new agencies and required them to operate to strict performance targets’ - this is interesting information, but to us it denotes activity, not output, and certainly not an outcome. The existence of the agencies and their targets in no way guarantees that the citizens and businesses using their services will have experienced any change in the level or quality of what they receive. It is significant that one of the best collections of comparative data on government reforms yet produced, the OECD’s *Government at a Glance* (2009) is principally concerned with (in our terms) inputs and processes. 159 pages of data do not contain a single table showing an input/output ratio (efficiency), still less a set of attributable outcomes (although the most recent, 2011 edition does tentatively begin some input/output analysis). Even when academics undertake extensive fieldwork in search of information about how reforms have affected performance, they may be unable to find it (Pollitt et al., 1998: 162-166). However, many academic articles describe how far techniques such as performance management or Total Quality Management have been installed in specified public departments or agencies (e.g. Joss and Kogan, 1995; Lægreid et al., 2008; Torres and Pina, 2004; Verheijen and Dobrolyubova, 2007). For us, once more, this is interesting, but not yet an output, still less an outcome. *If* the installation of TQM can be shown to have shortened public waiting times, then that *is* an output. And if the shortened waiting times and earlier treatment can be shown to have improved the outcomes for specific health problems (e.g. in cancer treatment or addiction problems) then those would be outcomes. However we should not assume that the mere articulation of a set of procedures (in this case TQM) will itself lead to such changes.

Thus we have taken a relatively relaxed approach to the process of admitting documents to our database, but a slightly puritanical approach to analyzing what it finally contains. We make no apology for this latter severity. Public management reform is a field extensively populated with individuals and organizations who have a stake in claiming successes. It is also a field, as we have seen, where the definition of many of the key terms is slippery and various. Our puritanism, if that is what it is, stems from a desire to achieve a measure of consistency across a large number of very diverse documents, together with our determination to focus, as far as we are able, on the outputs and outcomes that matter to the world *outside* public sector organizations. The question, to put it crudely, is ‘what do we *get* for all this reorganization?’ where ‘we’ are the citizenry and its representatives, civil society associations and companies. In all the multifarious literature on NPM this question has not been asked

often enough, and has rarely been asked on the empirical scale that is attempted here in COCOPS.

9. Key descriptive statistics from the database

Having now traveled through this dense thicket of definitional problems and methodological choices we can now begin to look at what has actually been found. According to the selection criteria described above, we ended up with a grand total of *510 documents*. These were distributed between the different countries and the European Commission as follows:

Table 2. Distribution of studies by country

Country	Number of studies*	Percentage**
Austria (AT)	28	5.5
Belgium (BE)	38	7.5
Croatia (HR)	22	4.3
Czech Republic (CZ)	25	4.9
Denmark (DK)	17	3.3
Estonia (ES)	23	4.5
Finland (FI)	14	2.8
France (FR)	70	13.7
Germany (DE)	63	12.4
Greece (GR)	11	2.2
Hungary (HU)	46	9.0
Ireland (IE)	12	2.4
Italy (IT)	58	11.4
Latvia (LV)	10	2.0
Lithuania (LT)	17	3.3
Luxembourg (LU)	9	1.8
Netherlands (NL)	63	12.4
Norway (NO)	37	7.3
Poland (PL)	12	2.4
Portugal (PT)	19	3.7
Romania (RO)	21	4.1
Slovakia (SK)	20	3.9
Spain (ES)	23	4.5
Sweden (SE)	23	4.5
Switzerland (CH)	13	2.6
United Kingdom (UK)	95	18.6
European Commission (EC)	8	1.6

* Includes both single-country and comparative studies. To avoid double counting the studies should not be added across countries. ** Percentage of the total database of 510 studies.

There are some evident differences in the number of studies included for various countries with five of them, the UK, Germany, the Netherlands, France and Italy, being better represented (each more than 10% of the total studies) than the rest. Overall, however, for most states (18 out of 27) the difference in the numbers is small, each of them with fewer than 5% of the total database. The figures include both single country and comparative studies, with varying degrees of treatment; therefore simply because a country had more studies included in the database does not automatically mean that it received a more intense treatment. Table 2 is complementary to table 1 in which we reported the degree of treatment for each country.

Tables 3 to 6 below describe the database with respect to sectors, type of documents, methods used and particular NPM-type reforms, tools and mechanisms identified.

The database suggests that the extent of evaluation of NPM reforms varies across sectors. Some sectors, such as education and health, appear to have been subjected to more evaluation than others, such as defence or spatial and urban planning. This may also reflect different degrees to which NPM-type reforms have been implemented across sectors, with health and education in the lead. One can also notice a sizable number of studies that evaluated parts of central government such as the core executive, the central government or the public sector more generally. Although our emphasis was on central government reforms, a high number of local level studies were eventually included, as shown in table 3.

Table 3. Distribution of studies by sector

Sector	Number of studies*	Percentage**
Core executive (excluding line ministries)	126	24.7
Legislature and associated bodies	17	3.3
Employment services	57	11.2
Health care, education and social services, except employment	161	31.6
Business and economic policy	60	11.8
Defence	21	4.1
Justice, law and order	50	9.8
Spatial and urban planning	20	3.9
Local services/local government	110	21.6

Public sector more generally, including central government	148	29.0
Other	25	4.9

* Due to many multiple-sector studies, the columns should not be added to avoid double counting.

** Based on the total database of 510 studies.

Table 4. Distribution of studies by type

Type of study	Number of studies	Percentage*
Academic	347	68.0
General official policy reports with some elements of evaluation	38	7.5
Internal formal official evaluation studies	28	5.5
External and independent official evaluation studies	54	10.6
Management consultancy reports	18	3.5
Studies by international or supranational bodies	12	2.4
Studies by civil society organizations	6	1.2
Other	7	1.4

* Based on the total database of 510 studies.

Academic studies by far outnumber any of the other type of study. However, more than 30% of the total database consists of non-academic studies (table 4). As is the case with reviewing studies using a variety of methods (table 5) we think that going beyond conventional approaches, i.e., include only academic studies using one type of method, can add data and value to our analysis and conclusions. In this sense our meta-analysis differs from other major meta-analyses in public management more generally.

Table 5. Distribution of studies by methods

Methods	Number of studies*	Percentage**
Mainly declarative (limited evidence)	92	18.0
Broad synthetic overview (there is an analytic attempt to make an assessment)	218	42.8
Single case study	66	12.9
Multiple case studies	119	23.3
Hypothesis/hypotheses testing	17	3.3
Historical descriptive narrative	46	9.0
Experimental method	4	0.8

Mainly quantitative analysis, but not experimental	97	19.0
--	----	------

* Due to many multiple-method studies, the columns should not be added to avoid double counting.

** Based on the total database of 510 studies.

Table 6. Distribution of studies by NPM reforms, tools and mechanisms

NPM tools and mechanisms	Number of studies*	Percentage**
Corporatization	36	7.1
Market-type mechanisms more generally (MTMs)	114	22.4
Contracting out	65	12.8
Purchaser-provider split	15	2.9
Performance league tables	26	5.1
Performance-related budgeting	40	7.8
Performance-related pay	37	7.3
User-choice mechanisms	20	3.9
Other MTMs	22	4.3
Performance measurement and management systems more generally	208	40.8
Normalization of public sector employment	39	7.7
Employment of management consultants	9	1.8
Use of internal contractual/quasi-contractual frameworks for steering	80	15.7
Creation of semi-autonomous agencies	63	12.4
Public service quality improvement schemes	53	10.4

* Due to many multiple-tool/mechanism studies, the columns should not be added to avoid double counting.

** Based on the total database of 510 studies.

Table 6 above suggests that the various elements of NPM have not been evaluated to the same extent. Some reforms, such as performance management and measurement systems or contracting out or market-type mechanisms appear in the literature more often than other type of reforms. The categories are designed to group together (and have grouped together) broadly similar tools and mechanisms.

10. A preliminary interpretation of the findings

The different analyses of the database relate to different subsets of the database contents (which subset is in play is always specified). In some cases we have analysed only those studies which include findings about outputs and outcomes. In others we have also included studies which focus on changes in activities/processes and say little or nothing about outputs

and outcomes. Table 7 (below) gives some idea of the major subsets we have used, in descending order of size.

Table 7. Some subsets of the COCOPS SharePoint database

Subset	Number of studies	Percentage*
1. All studies (total database)	510	100.0
2. All studies with at least one entry for empirical impacts	405	79.4
3. All studies including entries for changes in activities/processes, outputs or outcomes	354	69.4
4. All studies including entries for changes in outputs or outcomes	156	30.6
5. All studies with entries for changes in outputs or outcomes in central government only	131	25.7
6. All studies with entries for changes in outcomes	56	11.0
7. All studies with entries for changes in outcomes in central government only	48	9.4
8. All studies with entries for changes in quality	53	10.4
9. All studies with entries for changes in social cohesion	121	23.7

* Percentage of the total database of 510 studies.

One could say that the further one goes down the list from 1 to 7, the closer one comes to our ideal aim - to establish what actual final impacts (outcomes) NPM reforms in central government have produced. Rows 8 and 9 (quality and social cohesion) are somewhat separate and complex dimensions, as indicated by the discussion previously in this paper.

A total of 510 studies and reports of various types have been reviewed and coded according to an analytical framework. In a little more than four hundred of them at least one type of impact (generously construed) has been identified. Many of the studies we reviewed are complex and contain various types of 'impacts', from, say, transparency to staff motivation and morale to outcomes. Therefore one should not conclude that of the 405 studies (subset 2), 354 have commented on effects on outcomes, outputs or processes while only 51 (the difference between subset 2 and subset 3) have commented on effects other than outcomes, outputs, processes or activities.

One clear conclusion from table 7 is that many of the studies have not gone beyond reporting changes (or no change) in processes or activities. Effects on outputs are less common than changes in processes and they typically include effects on efficiency and productivity, quality of services and quantity of outputs. However, these terms are loosely defined and if we adopt a more conservative approach to defining them, we end up with even fewer studies. As for outcomes, we found that only 56 studies of the total database or a little more than one out of ten have gone that far. A typical outcomes-related measure is effectiveness, or the degree to which outcomes meet stated objectives of the reforms. One can adopt an even more conservative definition of effectiveness by including only measures of effectiveness that pertain to effects produced in the real, outside world of users and citizens or businesses. In most of the 56 studies this has been the case, but in some cases we allowed a more generous definition of effectiveness and treated it as an outcome in itself by its virtue of reflecting goals that have been accomplished.

Subtracting the studies that evaluated outcomes of local services, we ended up with only 48 or 9.4% of the total database with at least an entry for outcomes in central government.

Finally, 53 studies in the database had at least an entry for quality, mostly quality of services (outputs-type quality) or quality in operations, processes or activities. Based on our categorization of social cohesion in terms of effects on equal access to services and on civil servants' motivation and morale, we found that 121 or 23.7% of the studies included at least some comment on social cohesion.

Below in tables 8 and 9 we report some additional, preliminary findings. The data indicate that overall in broad terms, the impacts of NPM reforms have been assessed as more positive than negative. Outcomes and outputs, not just processes and activities, are found to have improved more often than not. One should note, however, that this broadly favourable picture is by no means uniformly positive. As table 8 shows, we found many studies with some evidence or comments on negative changes in outcomes, outputs or processes. A sizable number of studies and entries have not reported significant changes (or have reported no changes at all). For instance, more than 30% of the total entries for outcomes are of this kind.

Table 8. Summary of impacts of NPM-type reforms across Europe

Direction of 'change'	Up/Improved		Down/Worse		Unchanged		Uncertain	
	Number of entries	%*	Number of entries	%	Number of entries	%	Number of entries	%
Extent of impacts								

Outcomes	28	43.8	15	23.4	18	28.1	3	4.7
Outputs	90	60.4	27	18.1	27	18.1	5	3.4
Processes/ activities	360	59.4	111	18.3	120	19.8	15	2.5

* For each type of impact the percentages are based on the total number of entries for that particular type of impact.

A key point to note when interpreting the findings is that the labels we used, such as ‘up’ or ‘improved’ or ‘worse’ can carry either qualitative or quantitative connotations, or both. This is particularly so for effects on processes/activities and to some extent for outputs, but less so for outcomes which, by their very essence, are qualitative (‘soft’ impacts) *par excellence*. Up to this point we have not distinguished between qualitative improvements in processes and ‘simple’ changes in processes, such as the introduction of a performance management system or a public service quality improvement scheme. The latter may or may not lead to positive improvements in the quality of the process, and therefore improvements in outputs and possibly outcomes. Most of the 360 entries for positive changes in processes reflect quantitative changes in systems, processes and activities (i.e. they record that something new has been introduced) and only a minority reflects documented qualitative changes (i.e. the process has actually improved, against some relevant normative standard).

What is the story behind the recorded impacts on outputs and outcomes? Table 9 below tentatively presents some more specific findings on the outputs identified. Additional, information on more specific findings on outcomes would be too slippery to report at this point therefore we decided to leave this for a later stage of our research. The data on outputs suggest that:

- More studies have assessed the effects of NPM-type reforms on efficiency and productivity than on changes in quality and quantity of outputs (61 entries compared to 40 and 29 respectively)
- Measures related to outputs regardless of their type are found to have improved in more cases than they were found to have decreased or remained unchanged
- This favourable picture is particularly evident in the case of efficiency and productivity and less so for quality of services and for quantity of outputs
- A relatively significant number of ‘unchanged’ entries on quality (25%) were found which shows the complex nature of assessing changes in the quality of public services.

Table 9. Findings on outputs

Direction of 'change' 'Type' of outputs	Up/Improved		Down/Worse		Unchanged		Uncertain	
	Number of entries	%*	Number of entries	%	Number of entries	%	Number of entries	%
Quantity of outputs	18	62.1	10	34.5	1	3.4	0	0.0
Quality of services and outputs	19	47.5	8	20.0	10	25.0	3	7.5
Efficiency and productivity	41	67.2	7	11.5	9	14.8	4	6.6
Other	12	57.1	2	9.5	5	23.8	2	9.5

* For each type of outputs the percentages are based on the total number of entries for that particular category.

Finally, it is important to make the basic point that improvements in one aspect may be balanced out by deteriorations in another. Thus, to say that the majority of those studies which focus on outputs record improvements does not mean that there may not also be deteriorations recorded in some of those studies as well. For example, efficiency may improve but equity of access declines. Or quality of service is improved but only at the cost of higher input expenditure.

11. Implications for further research

This paper is the first in a series, and only scratches the surface of what can be found in the database. Whilst we have here offered the big, surface picture, the database will permit ourselves and other scholars to conduct more fine-grained analyses of many aspects. These could include, for example, differences between countries or sectors, differences over time (is the pattern of impacts from the 1980s and early '90s significantly different from those since 2000?) or a close analysis of the contextual factors which are mentioned in many studies, and which may yield patterns. One unusual feature of this database is that it is exclusively concerned with NPM in the EU, whereas a large slice of the world literature on NPM-type reforms concerns the US, Australia, Canada and New Zealand. There is therefore a wider comparative opportunity for scholars to use this database as one leg of a comparison with other countries or regions outside the EU.

12. Concluding observations

What, finally, can we say about the impacts of NPM in Europe? The situation appears paradoxical. On the one hand there have been endless publications - both academic and official - concerned with NPM-like programmes and techniques. Yet, on the other, our solid, scientific knowledge of the general *outcomes* of all this thinking and activity is very limited.

Claims and counterclaims outnumber hard, carefully collected evidence by a substantial margin. That was the case in the mid 1990s (Pollitt, 1995) and it remains the case today. A summarising graphic metaphor might be that there is an ocean of studies of the application of NPM ideas within the EU, but only a modest sea of works that offer direct empirical analysis of *outputs*, and no more than a small pond that convincingly connect specific reforms to particular *outcomes*.

It is true that our database holds a good deal of interesting information about changes in *outputs* resulting from NPM, but what the collectivity of this knowledge suggests is that these vary enormously, from highly positive through 'no change' to rather negative, depending on circumstances. And the circumstances seem to be very various. Further exploration of these contextual influences will be one important focus for later working papers from the COCOPS team. However, the majority of our publications, however, get no further than discussions and analyses of changes in *processes and activities*. That is the 'ocean'. Thus, all generalisations of the form 'NPM reforms lead to outcome X' are suspect, at least in the sense that the mountain of evidence in our database will not yield any such firm and overarching conclusion.

If this is the picture we arrive at after a meta-analysis of hundreds of documents ostensibly concerned with the impacts of NPM, what conclusions can be drawn? Broadly speaking, there are two obvious possibilities. First, perhaps there *are* general rules which adequately summarize the impact of NPM reforms, but we have not found them. Alternatively, second, there may be no such general rules, in which case we may have been asking the wrong kind of question.

Let us consider the first alternative first. Why might our research have failed to find the systematic connections between NPM reforms (reform 'inputs' if you will) and the outputs and outcomes of activities conducted by public organizations? Here there are at least two possibilities. One is that our literature search was incomplete, and that there are documents out there, somewhere, that would yield more definite general connections and rules. This is a possibility. The literature is huge and very diverse, and it is always possible that we and our partners and collaborators have overlooked some crucial analysis. If so, we hope that those who know of such missing links will let us know. On the whole, though, we find the probability of the existence of some missing, yet definitive analysis - something that will provide a *general* rule or formula - fairly low. [That we have missed *something* is almost certain, but that we have missed something *crucial* is unlikely.] Another possibility is that useful knowledge exists, but that it is hard – or even impossible – to put into a codified,

explicit, 'scientific' form. It may be a form of 'craft' or tacit knowledge, which experienced practitioners have developed but which depends on unarticulated understandings of the nuances of particular situations and formations of reform actors. In fact some writers have suggested that this type of knowledge is of great (but largely unacknowledged) importance in organizational practice (Tsoukas, 2005). If, however, this is the case then, *ipso facto*, we are not going to find the general rules by searching academic and official literature. One would have to adopt quite different research strategies, for example by using prolonged participant observation of experienced public managers at work.

The second main possibility is that there are no general rules to be had. It is vital to understand that this does *not* mean that we can have no knowledge of NPM reform impacts. It simply means that the knowledge is unlikely to come in the form of a general, 'Z follows Y follows X' rule. Perhaps the topic is just too complex and contingent on many varying factors and influences for that kind of stable general relationship to be available. However, more specific, context-dependent or *ad hoc* forms of knowledge may well exist. And that, we would argue, is one of the things that our meta analysis points towards. In this respect the findings about contexts are crucial. It is by carefully sifting those that more particular and conditional (and local) sets of relationships may be identified.

Already, some context-dependent generalisations suggest themselves from the literature in the database. For example:

- Performance related pay requires a long list of supportive local conditions before it stands a good chance of working as intended. In particular it tends not to work well where a) political patronage determines most senior appointments, b) the bonuses available are only a very small % of the total remuneration, c) the work is hard to measure in an objective and widely-accepted way (and so on)
- Performance targets can definitely have a significant impact on performance, but usually only where backed up by significant penalties/incentives
- Contracting out often fails to yield significant savings, particularly in the medium and long term. Again, the context has to be right for contracting out to work in the way it is supposed to. Necessary supports include a) the possession of contract design and management skills by the staff of the parent public sector organization, b) the presence of real competition (which may exist at first but then disappear because of private sector mergers and take-overs) and c) that the activity being contracted is one which can be specified in fairly firm detail

- Use of market-type mechanisms (MTMs) may work better in simpler, more standardized services than in complex, unstandardized, professionalised human services. There are a number of studies in a number of countries indicating problems with the application of MTMs to healthcare and educational services
- The politico-administrative culture is mentioned as a significant factor shaping reform impacts in many studies – especially in Eastern and Central Europe, but also in the West.

At present these connections and relationships are only impressionistic. One of our next tasks is to interrogate the database in detail to see how far these, and other, contextual relationships can be confirmed and refined.

Our next research paper will therefore deal with the issue of context in greater depth.

Bibliography

- Andrews, R., Boyne, G. & R. (2011) 'The impact of management on administrative and survey measures of organizational performance', *Public Management Review* 13(2), 227-255.
- Audit Commission (2009). *Final score: the impact of Comprehensive Performance Assessment of local government, 2002-2008*. London: The Audit Commission.
- Boyne, G., Farrell, C., Powell, M. & Walker, R. (2003). *Evaluating public management reforms*. Buckingham: Open University Press.
- CEEP (2010). *Public services within the European Union and the 27 member states: statistics, organizations and regulations*. Paris: project with support of the European Commission, May.
- Chan, J., Ho-Pong, T. & Chan, E. (2006). Reconsidering social cohesion: developing a definition and analytical framework for empirical research. *Social Indicators Research*, 75(2), 273-302.
- Cooper, H. (1998). *Synthesizing research*. Thousand Oaks: Sage.
- Clifton, J. & Díaz-Fuentes, D. (2010). Evaluating EU policies on public services: a citizens' perspective'. *Annals of Public and Cooperative Economics*, 81(2), 281-311.
- Dorrell, S. (1993). *Public sector change is a world-wide movement*. Speech by the Financial Secretary to the Treasury, Stephen Dorrell, to the Chartered Institute of Public Finance and Accountancy, London, 23 September.
- Dunleavy, P., Margetts, H., Bastow, S. & Tinkler, J. (2006a). Digital-era governance: IT corporations, the state and e-government. Oxford: Oxford University Press.
- Dunleavy, P., Margetts, H., Bastow, S. & Tinkler, J. (2006b). New Public Management is dead – long live digital-era governance. *Journal of Public Administration Research and Theory*, 16(3), 467-494.
- Federal Ministry of the Interior (2006). *Focused on the future: innovations for administration*. Berlin: Federal Ministry of the Interior.
- Ferlie, E. & Geraghty, J. (2005). Professionals in public service organizations: implications for public sector reforming. In E. Ferlie, L. Lynn Jr. & C. Pollitt (Eds), *The Oxford handbook of public management* [pp. 422-445]. Oxford: Oxford University Press.
- Gore, A. (1997). *Businesslike government: lessons learned from America's best companies*. Washington, DC: National Performance Review.
- Gow, J. & Dufour, C. (2000). Is the New Public Management a paradigm? Does it matter? *International Review of Administrative Sciences*, 66(4), 573-593.
- Hood, C. (1990). De-Sir Humphreyfying the Westminster Model of bureaucracy: a new style of governance? *Governance*, 3(2), 205-214.
- Hood, C. (1991). A public management for all seasons. *Public Administration*, 69(1), 3-19.
- Hood, C. & Peters, G. (2004). The middle aging of New Public Management: into the age of paradox? *Journal of Public Administration Research and Theory*, 14(3), 267-282.
- Jacobs, R. & Goddard, M. (2007). How Do Performance Indicators Add Up? An Examination of Composite Indicators in Public Services. *Public Money & Management*, 27(2), 103-110.
- James, O. (2010). *End user satisfaction as a performance measure for public services*. Paper for the National Audit Office, Department of Politics, University of Exeter, February.
- Joss, R. & Kogan, M. (1995). *Advancing quality: Total Quality Management in the National Health Service*. Buckingham: Open University Press.
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2007). Governance matters VI: governance indicators for 1996-2006. World Bank Policy Research Working Paper 4280.

- Kearns, A. & Forrest, R. (2001). Social cohesion, social capital and the neighbourhood. *Urban Studies*, 38(12), 2125-2143.
- Kelman, S. & Friedman, J. (2009). Performance improvement and performance dysfunction: an empirical examination of the distortionary impacts of the emergency room wait-time target in the English National Health Service. *Journal of Public Administration Research and Theory*, 19(4), 917-946.
- Kurki, M. (2008). *Causation in international relations; reclaiming causal analysis*. Cambridge: Cambridge University Press.
- Lægneid, P., Roness, P. & Rubecksen, K. (2008). Performance information and performance steering: integrated system or loose coupling? In W. van Dooren & S. van de Walle (Eds), *Performance information in the public sector: how it is used*. Basingstoke: Palgrave.
- Lynn, L. Jnr. (2006) *Public management: old and new*, London and New York, Routledge/Taylor and Francis
- McNabb, D. (2010). *Case research in public management*. Armonk, NY: M.E. Sharpe.
- OECD (1995). *Governance in transition: public management reforms in OECD countries*. Paris: PUMA/OECD.
- OECD (2009). *Government at a glance 2009*. Paris: OECD.
- OECD (2011). *Government at a glance 2011*. Paris: OECD.
- Office of Public Services Reform (2002). *Better government services: executive agencies in the 21st century*. (<http://www.civilservice.gov.uk/agencies>).
- Ongaro, E. (2009) *Public management reform and modernization: trajectories of administrative change in Italy, France, Greece, Portugal and Spain*, Cheltenham and Northampton, Mass., Edward Elgar
- Osborne, D. & Gaebler, T. (1992). *Reinventing government: how the entrepreneurial spirit is transforming the public sector*. New York, NY: Addison Wesley Publishing Company.
- PA Consulting Group (2002). *Evaluation of the Strategic Management Initiative*. Report presented to the Department of the Taoiseach, Dublin, PA Knowledge Ltd.
- Painter, M. and Peters, G.B. (eds.) (2010) *Traditions and public administration*, Basingstoke, Palgrave/Macmillan
- Pawson, R. (2002). Evidence-based policy: the promise of “realist synthesis”. *Evaluation*, 8(3), 340-358.
- Pawson, R., Greenhalgh, T., Harvey, G. & Walshe, K. (2005). Realist review - a new method of systematic review designed for complex policy interventions. *Journal of Health Service Research Policy*, 10(11), S1.21-S1.34.
- Pollitt, C. (1995). Justification by works or by faith? Evaluating the New Public. *Evaluation*, 1(2), 133-154.
- Pollitt, C. & Bouckaert, G. (Eds) (1995). *Quality improvement in public services: concepts, cases and commentary*. London: Sage.
- Pollitt, C., Birchall, J. & Putman, K. (1998). *Decentralising public service management*. Basingstoke: Macmillan.
- Pollitt, C. & Bouckaert, G. (2000). *Public management reform: a comparative analysis*. Oxford: Oxford University Press.
- Pollitt, C., Talbot, C., Caulfield, J. & Smullen, A. (2004). *Agencies: how governments do things through semi-autonomous organizations*, Basingstoke, Palgrave/Macmillan
- Pollitt, C. (2003). *The essential public manager*. Buckingham: Open University Press/McGraw Hill.
- Pollitt, C. & Bouckaert, G. (2009). *Continuity and change in public policy and management*. Cheltenham: Edward Elgar.
- Pollitt, C. & Bouckaert, G. (2011). *Public management reform: a comparative analysis: New Public management, Governance and the Neo-Weberian State*. Oxford: Oxford University Press.
- Pollitt, C. (2011). Moderation in all things: international comparisons of governance quality. *Financial Accountability and Management* (in press).
- Prime Minister and the Minister for the Cabinet Office (1999). *Modernising government*, Cm. 413, London, The Stationery Office.

- Propper, C., Sutton, M., Whitnall, C. & Windmeijer, F. (2008). *Incentives and targets in hospital care: evidence from a natural experiment*. Bath, CMPO Working Paper 9.
- Roberts, A. (2010). *The logic of discipline; global capitalism and the architecture of government*. Oxford: Oxford University Press.
- Torres, L. & Pina, V. (2004). Reshaping public administration: the Spanish experience compared to the UK. *Public Administration* 82(2), 445-464.
- Torres, L., Pina, V. & Royo, S. (2005). E-government and the transformation of public administrations in EU countries: beyond NPM or just a second wave? *Online Information Review*, 29(5), 531-553.
- Tsoukas, H. (2005). *Complex knowledge: studies in organizational epistemology*. Oxford: Oxford University Press.
- Van de Walle, S., Van Roosbroek, S. & Bouckaert, G. (2008). Trust in the public sector: is there any evidence for a long term decline? *International Review of Administrative Sciences*, 74(1), 47-64.
- Verheijen, T. & Dobrolyubova, Y. (2007). Performance management in the Baltic states and Russia: success against the odds? *International Review of Administrative Sciences*, 73(2), 205-215.

Appendix A: List of collaborators

Dan Apăteanu, Babeş-Bolyai University, Romania
Yuliyana Atanasov Gospodinov, Tsenov Academy of Economics, Bulgaria
Maria Asensio, National Institute of Administration, Portugal
Dario Barbieri, Bocconi University, Italy
Dana Bajusova, Economics University of Bratislava, Slovakia
Ioana Ciucanu, National School of Political Studies and Public Administration, Romania
Anna Czarczynska, Leona Kozminskiego Academy Warsaw, Poland
Iwona Grabowska, Collegium of World Economy, Warsaw School of Economics, Poland
Tomas Jacko, Comenius University Bratislava, Slovakia
Milena Krasimirova Vladimirova, Medical University of Varna, Bulgaria
Ivan Koprivic, Zagreb Law Faculty, Croatia
Georgi Dimireov Manliev, Technical University of Sofia, Bulgaria
Monica Marin, Research Institute for Quality of Life, Romania
Lucica Matei, National School of Political Studies and Public Administration, Romania
Miruna Mazurencu-Marinescu, Academy of Economic Studies, Romania
Marco Meneguzzo, University of Rome Tor Vergata, Italy
Michał Mierzwa, Krajowa School of Public Administration Warsaw, Poland
Beata Mikusova Merickova, Matej Bel University Banská Bystrica, Slovakia
Cristina Mititelu, University of Rome Tor Vergata, Italy
Vitalis Nakrošis, Vilnius University, Lithuania
Juraj Nemeč, Matej Bel University Banská Bystrica, Slovakia
Goranka Lalic Novak, Social Sciences Polytechnic Zagreb, Croatia
David Spacek, Masaryk University, Czech Republic
Robert Szczepankowski, Stanisław Staszic College of Public Administration, Białystok,
Poland
Piotr Szreniawski, Maria Curie Skłodowska University, Lublin, Poland
Ioan Toderaş, Centre for Promoting Research and Innovation, Romania
Adelina Dumitrescu, National School of Political Studies and Public Administration,
Romania
Agne Vabamäe, Tallin University of Technology, Estonia
Mădălina Voican-Avram, University of Craiova, Romania
Dominika Wojtowicz, Leona Kozminskiego Academy Warsaw, Poland
Margarita Yordanova Bogdanova, Tsenov Academy of Economics, Bulgaria

COCOPS Working Papers

1. Coordinating for Cohesion in the Public Sector of the Future: COCOPS Project Background paper
2. The Brave New World of Public Infrastructure: Is Market-Oriented Reform Producing a “Two-Track” Europe?
3. The Impacts of the New Public Management in Europe: A Meta-Analysis

