Implementing Public Accounts Committee Recommendations: Evidence from the UK Government's 'Progress Reports' since 2012

Thomas Elston (iD) [1],* **and Yuxi Zhang** (iD) [2]

[1]*Blavatnik School of Government, University of Oxford, Oxford, OX2 6GG, UK;*
[2]*School of Public Policy, London School of Economics and Political Science, London, WC2A 2AE, UK*

*Correspondence: thomas.elston@bsg.ox.ac.uk

The Public Accounts Committee is the oldest and, reputedly, most influential oversight committee in the British parliament. Since 2012, the government has published an unprecedented volume of data explaining its actions in response to the PAC's inquiries. By tracking all 615 of the Committee's recommendations from 2010 to 2012 across the first 14 of these 'progress reports', we explore how effectively this novel reporting mechanism serves to inform parliament about the executive's responsiveness to scrutiny. We find the reports to yield rich data on the fate of many individual recommendations, but only limited information on the overall timeliness and quality of implementation. Flexibilities and inconsistencies in reporting style also make the arrangements vulnerable to gaming. We identify potential improvements and research opportunities.

Keywords: Accountability, Audit, Bureaucracy, Civil service, Oversight, Select committees

I beg to move: That this House calls on the Government to ensure that all recommendations contained in Reports of the Committee of Public Accounts and accepted by Government Departments are implemented, and that the relevant Minister makes a statement to the House on any recommendations accepted but not implemented within a year of their acceptance.

—Margaret Hodge MP, December 2010
(Hansard, Vol. 520, Col. 117)

1. Introduction

The Public Accounts Committee is the oldest and, reputedly, most influential oversight committee in the British parliament. Charged with scrutinising probity and value for money in public expenditure across the whole breadth of government, the PAC is supported by the Comptroller and Auditor General (a similarly long-lived office) and an 800-strong staff at the National Audit Office (the UK's 'supreme audit institution'). The PAC's reputation is formidable, referred to by Hennessy as 'queen of the select committees' (Public Accounts Committee [PAC], 2007, p. 7) and by Staddon (2015) as 'first among equals'. As enforcer of the civil service accounting officer system, the Committee is especially credited with 'keeping many senior mandarins on their toes' (Rogers *et al.*, 2019, p. 267). Moreover, its inquiries, like those of the NAO, frequently make news headlines, with broadcasters and columnists habitually declaring that: 'An influential group of MPs has found ...'

But gauging the influence of legislative overseers on public policy is fraught with difficulties, particularly when such committees have limited formal powers (Hindmoor *et al.*, 2009; Benton and Russell, 2013; Gaines *et al.*, 2019). Similarly, since the scope of contemporary audit is vast (Power, 1997), we lack counterfactuals from which to infer its unique impact on auditees. The mere prospect of investigation in future may alter decision-making in the present—the so-called 'psychological deterrent' (Flesher and Zarzeski, 2002, p. 94). And of the many potential channels to exert influence over government, those that are most straightforwardly measured and attributed—such as the formal acceptance of inquiry recommendations—reveal relatively little about the practical impact of oversight (Morin, 2008; Kells, 2011). Conversely, more informative metrics—like whether recommendations are ever actually *implemented*, how quickly and thoroughly, and with what cumulative effect on behaviour or outcomes—present many difficulties in measurement and inference (Hindmoor *et al.*, 2009; Benton and Russell, 2013).

Nonetheless, it remains a priority to better understand executive-legislature accountability, its impact (good or otherwise) on political and bureaucratic behaviour, and its potential to drive improvements in public services and trust in political institutions (Dubnick, 2005; Bovens *et al.*, 2008; Keane, 2009; Flinders, 2012). Furthermore, with public expenditure under pressure and growing appreciation of the potential downsides of performance auditing, such as encouraging risk aversion and ritualistic compliance (Kells, 2011), overseers feel compelled to demonstrate their own 'value added' (Talbot and Wiggan, 2010). The NAO's claimed return 'for the nation', most recently valued at '£16 of positive financial impact for every £1 the NAO costs' (NAO, 2020, p. 33), is a case in point.

In this context, a little-noticed reform of a decade ago is intriguing. In January 2012, the Treasury published the first of what was to become (so far) 16 'progress reports' explaining steps taken to fulfil each government-accepted PAC recommendation. This extensive documentation, which now amounts to more than 3000 pages, offers new and potentially rich evidence on government *actions*—not simply *declared intentions*—following parliamentary scrutiny.

Drawing on a novel dataset that tracks all 615 PAC recommendations from the 2010 to 2012 parliament across the first 14 of these progress reports, this article provides an initial examination of the new reporting device. Our aim is, firstly, to understand how effectively it updates parliament on the speed and thoroughness of the executive's response to scrutiny. We then consider how the arrangements might be improved and extended, and what new research questions may become answerable through the data.

We find the progress reports to provide rich detail on individual cases of committee intervention, and a basis for forming a broader assessment of executive responsiveness. We show, for instance, that two-thirds of recommendations are achieved within 12 months of the timeline initially set out by government; and that the nature or 'task' of the recommendation may influence the pace of its implementation. However, the overall value of the reporting system is impeded by the irregular publication schedule, inconsistencies in how departments disclose data, and multiple vulnerabilities to gaming by any organisation minded to exaggerate its compliance with the committee's wishes. Small improvements in reporting design and enforcement could yield significant gains to accountability. Moreover, the data also raise the prospect of undertaking valuable new research into factors affecting the speed and quality of government action in response to oversight, the alleged strength of the PAC compared with other select committees, and whether this is dynamic over time.

Below, the second section introduces the PAC, NAO and wider select committee system, and the third explains the reporting reforms of 2012 and derives criteria for judging their effectiveness. The fourth section explains the construction of the dataset and the fifth presents the analysis. The sixth evaluates the overall effectiveness of the monitoring arrangements and recommends improvements, and the seventh outlines possibilities for future research.

2. The Committee of Public Accounts

2.1 Historical origins

The House of Commons Committee of Public Accounts, generally known as the Public Accounts Committee or simply the PAC, dates from 1861—when Gladstone was chancellor, public spending stood at less than 8 per cent of GDP,

and the civil service was experiencing the long hiatus between the announcement and eventual realisation of the Northcote-Trevelyan reforms (Hennessy, 1990). The Committee was tasked with examining the government's financial accounts, 'showing the appropriation of sums granted by Parliament to meet the public expenditure' (PAC, 2007, p. 4). This charge it retains today, amid the huge growth in the size and complexity of the state since the Victorian age. In practice, however, the PAC's remit has evolved, not least because the mandate and powers of its principal informant, the National Audit Office, has similarly expanded (Staddon, 2015).

Five years after the PAC's inception, in 1866, government legislated for a new statutory officeholder, the Comptroller and Auditor General,[1] to support financial oversight (Dewar and Funnell, 2016). The Comptroller was to authorise the issuing of money to government departments and interrogate government accounts, aided by a new audit department within the civil service (White and Hollingsworth, 1999). This proximity to the executive was contentious from the outset, yet it took more than a century to achieve independence. The National Audit Act 1983 recreated the Comptroller as an 'Officer of the House of Commons', supported by a new National Audit Office (Dewar and Funnell, 2016). The Comptroller's statutory duties expanded to include discretionary investigations into economy, efficiency and effectiveness (Pollitt *et al.*, 1999), legitimating the so-called 'performance auditing' already being undertaken informally, and laying the foundation for the growth of the PAC's 'value for money' focus (Staddon, 2015). In undertaking this work, the NAO, like other 'supreme audit institutions', faces the difficult balancing act of producing reports that are useful for performance improvement yet sufficiently detached to aid robust scrutiny of the executive (on this dilemma, see Lonsdale, 2008).

Slightly earlier, in 1979, the House of Commons created a series of new select committees charged with overseeing policy, administration and expenditure in individual government departments (Drewry, 1985; Judge, 1992; Aylett, 2019). The immediate effect on the enduring PAC is unclear. Robinson (1985) notes that departmental committees generally avoided financial scrutiny of their charges (see also Procedure Committee, 2019); but potential overlap in, and competition between, their inquiries and those of the PAC has also been noted (Geddes, 2020, p. 134). In addition, subsequent to 1979, the PAC has been subject to the broader reforms designed to further develop select committees as a whole. For instance, like other committees, its chair is now remunerated (since 2003) and elected by the whole House (since 2010), and members are also elected by party groups (Russell, 2011). Further, instead of the day of 'set piece' debate

---

[1]The earliest reference to an official being charged with auditing government finances actually dates from 1314 (NAO, 2021).

reserved annually for the PAC, since 2010 the Committee has needed to bid for time in the main chamber from the new Backbench Business Committee alongside other committees and individual MPs (Foster, 2015).

2.2 Modus operandi

At 14 members, the PAC is slightly larger than most departmental committees (Russell and Gover, 2017, p. 207). It also meets more frequently, with two or occasionally three hearings a week (Hodge, 2016). Uniquely, the PAC Chair is always drawn from the opposition, and a Treasury Minister holds ex officio membership (but rarely attends) so that government receives committee papers directly (White and Hollingsworth, 1999). The remaining membership reflects the composition of the House, providing a government majority.

Like other committees, the PAC can summon witnesses and request documents; but, given its support from the Comptroller, it was unable to engage additional advisers until 2010. The monarch appoints the Comptroller upon recommendation of the House of Commons, which votes on a motion proposed by the Prime Minister after agreement with the PAC Chair (Budget Responsibility and National Audit Act 2011, S11). The Comptroller has discretion in selecting inquiry topics, though must 'have regard' to proposals from the PAC (National Audit Act, 1983, S7 as amended). Informally, the NAO also consults with auditees to ensure inquiry relevance and/or determine timing and methods (Lonsdale, 2008, p. 235). The 1983 Act (S8) provides the Comptroller with access to 'all such documents as he may reasonably require' and the NAO submits draft reports to auditees for fact checking (White and Hollingsworth, 1999; Lonsdale, 2008). Consequently, the point of departure for a PAC hearing is typically more developed than for other committees, exemplifying, as Brazier and Fox (2011, p. 364) suggest, 'how technical expertise can be combined with political input in the scrutiny process.'

The Committee is the focal point of the civil service accounting officer system, whereby senior bureaucrats are personally accountable for the regularity, propriety, value for money, and feasibility of public expenditure (HM Treasury, 2015). Of the 412 civil servants that gave evidence before the 24 select committees studied by Geddes (2018) in 2013–14, more than a quarter were attending the PAC.[2] Historically, these hearings were low-key affairs, described as 'quiet and informal' and requiring minimal 'weekend' preparation (Chapman, 1978, pp. 73, 119). However, Rhodes's (2011, p. 110) contemporary ethnography of permanent secretaries reveals this to have altered considerably, with 'intensive dress rehearsals' now proceeding PAC interrogations, which are televised, often brusque, and occasionally

---

[2]Analysis of data used in Geddes (2017) kindly shared by the author.

heated. These rehearsals constitute a 'PAC damage limitation exercise' (White and Hollingsworth, 1999, p. 123), with further defence provided in the Treasury's *Accounting Officer's Survival Guide*, which identifies strategies to 'reduce the risk of difficult PAC hearings' (HM Treasury, 2015, pp. 5–6).

Alongside accounting officers, the PAC summons 'senior responsible owners' of projects to give evidence (PAC, 2015), as well as civil servants who have changed position since the events under investigation occurred (Hodge, 2016). During 2010–15, and reflecting the Committee's preoccupation with tax evasion at this time, high-profile hearings were also held with figures from big business, including Google, Amazon and Starbucks. Conversely, only one minister appeared before the Committee over this whole period (Hodge, 2016, p. 22).

2.3 Sources of influence

A combination of factors explains the formidable reputation of the PAC among commentators; although there is limited research testing this viewpoint empirically.

Firstly, this orthodoxy may speak more of the perceived weakness of the departmental committees against which the PAC is compared than of the observable effect of the PAC itself. Unlike in legislatures with 'strong' committee systems, the 1979 reforms gave select committees no role in either budget formation or legislative initiation (both still (almost) exclusively the preserve of the executive), nor in the scrutiny of bills (which continued in ad hoc committees). This led many to question the impact of departmental committees (Judge, 1992; Mattson and Strøm, 1995; Powell, 2000), although gradual reform and reassessment have recently led some to challenge this stereotype (Benton and Russell, 2013; Russell and Gover, 2017; Geddes, 2020; Matthews, 2020).

Secondly, the wide and unfettered focus of PAC inquiries, coupled with the personalised responsibility of accounting officers, is suggested to impose a kind of 'automatic check' on decision-making among officials. Hennessey opines that, 'by its very existence [the PAC] exert[s] a cleansing effect in all government departments' (PAC, 2007, p. 7). But how and in what conditions this deterrent manifests, and to what degree it extends outside of departments to decentralised and outsourced public services, remains uncertain. It is also unclear that this effect is unique to the PAC among other oversight committees (see Benton and Russell, 2013, p. 792).

Thirdly, since the authorising of tax and spending—the 'power of the purse'— is a primary legislative control over the executive, the role of the PAC in ex-post financial scrutiny may increase its perceived status, especially in a context of weak ex-ante controls (see Procedure Committee, 2019). Furthermore, the PAC's pursuit of value for money is expressly shared with the Treasury, which is sometimes

referred to as parliament's 'guardian' in Whitehall (Constitution Committee, 2013, pp. 9–12). This alignment of interests may create privileged opportunities for PAC influence at the heart of government (White and Hollingsworth, 1999); whereas, in respect of the departmental committees, the Treasury's fear in 1979 was that the reform would bring *increased* pressure for public expenditure (Aylett, 2019).

Fourthly, while departmental committees are free to examine policy decisions, the PAC's remit is more restricted. The seeming 'neutrality', even 'dryness', of audit and implementation may reduce partisanship among committee members, enabling greater consensus over report conclusions and recommendations (Hodge, 2016). This is often regarded as necessary to achieve influence with government (Geddes, 2020). The focus of recommendations on implementation rather than policy change (insofar as these are separable) may also allow greater buy-in from civil servants, since this (appears to) involve rethinking policy 'means', not 'ends'.

Fifthly, whereas low acceptance of committee recommendations is sometimes explained by the limited analysis that resource-stretched departmental committees can undertake in support of their ideas (Hindmoor *et al.*, 2009, p. 75), the PAC's work 'is made more credible by the very substantial back-up of the [NAO]' (Rogers *et al.*, 2019, p. 266). While the NAO has recently expanded its briefings for other committees (Midgley, 2019), the PAC retains the lion's share of its attention. Most PAC inquiries follow a prior NAO investigation; the Comptroller or a nominee attends all hearings; and NAO staff write the first draft of PAC reports (Staddon, 2015). (However, the reports of the NAO and PAC are not identical, and the NAO receives no formal response to its recommendations.)

Finally, amid growing news coverage of select committees generally (Kubala, 2011), research consistently ranks the PAC as either the first or second most-reported-upon committee by a large margin (Dunleavy and Muir, 2013; Gaines *et al.*, 2019). Its following on social media is also far greater (although the NAO's Twitter following is greater still). Such attention may in itself indicate a particular kind of societal influence (Gaines *et al.*, 2019); but it is also regarded as a valuable tool for gaining attention within government (Bringselius, 2014).

3. The executive's response

3.1 The development of progress reporting

Turning from the PAC itself to the executive's reaction to its investigations, government is obliged to respond to all select committee reports, normally within two months (Geddes, 2020). For the PAC, departmental responses are collated by the Treasury Officer of Accounts (an official in the finance ministry) and

published approximately bimonthly as 'Treasury Minutes'. Historically, these short acceptances or rebuttals constituted the government's 'final reply' to its overseers (Anand, 1988, p. 258). In the early 2000s, however, individual departments published limited information on the implementation of previously accepted PAC recommendations. Initially, this occurred in the spring or autumn performance reports that formed part of the Labour Government's expenditure planning and performance management system (Talbot, 2010). But latterly many departments appended the information to their financial statements each summer.

While this practice further distinguished the PAC from other committees, for whom no such implementation update was (or is) routinely provided, the decentralised approach led to much variation between departments. Some gave consistently detailed information, while others provided the bare minimum. Some only updated on the PAC's most recent inquires, while others dealt with all 'outstanding' matters not previously reported as complete.[3]

After the change of government in 2010, the new Chief Secretary to the Treasury swiftly cancelled the imminent spring reporting round and instructed departments to publish 'one table on outstanding PAC recommendations' in summer accounts. This was to be 'purely fact-based' with 'no narrative', these being recommendations accepted by the previous government and not necessarily endorsed by the new administration.[4] Then, in December 2010, in the first PAC debate arranged under the new Backbench Business Committee, the chair proposed a motion, cited in the epigraph above, calling on ministers to explain any accepted recommendations that remained unimplemented after 12 months. Government supported the motion, pending agreement on a suitable reporting format; and, in January 2012, the first 'progress report' was duly issued by the Treasury Officer of Accounts (not a minister, as the motion had proposed). So far, 15 further editions have appeared, most recently in November 2021.

The progress reports are prepared according to the following principles:

- Almost all PAC recommendations that are either fully or partly accepted are reportable (including those to arm's-length bodies.) The only exceptions are those immediately implemented at the time of acceptance, and those directed at fully independent bodies, like the BBC.

---

[3]The government strengthened its guidance on this for the 2009 reporting round, requiring that 'All outstanding recommendations made by the PAC must now be reported on, not just those made during the current reporting round' (see Public Expenditure System paper PES 2008 (17)).

[4]See PES 2010 (09).

- Updates are grouped according to their originating PAC inquiry. Mostly, updates are provided for each individual recommendation; but, occasionally, progress toward groups of related recommendations is reported together.
- Updates consist of a summary status classification ('implemented', 'work-in-progress', etc.) followed by a description of action taken or planned.
- Updates categorised as work-in-progress usually provide a 'target implementation date (TID)' of varying duration and precision. These may be revised once or more in subsequent updates.
- Recommendations are reported upon until self-classified as 'implemented' or, rarely, superseded by a new recommendation or policy change, or the government revokes its prior acceptance. (However, outstanding recommendations are not updated in every report, especially in the early editions, when the methodology was still being refined.)

(Online Appendix B contains an image of three pages of a progress report.)

3.2 Criteria for effective reporting

By what criteria should the success of these progress reports be judged? We focus on the following three yardsticks.

First is that the reports genuinely inform parliament about *actions*, not merely declared *intentions*. Scholars of select committees (Aldons, 2000; Hindmoor *et al.*, 2009; Benton and Russell, 2013) and performance auditing (Morin, 2001; Kells, 2011) both caution against regarding the acceptance of a recommendation as a guarantee of its implementation. Government may accept in good faith but later renege, perhaps as the difficulty or appropriateness of the action becomes clear or as other policy or political challenges and opportunities arise. Alternatively, acceptance may be symbolic, intended to reassure interest groups or placate a hostile committee or media without materially changing government policy or practice.

But what information on implementation would be useful and use-able? Elite politicians can experience information overload (Walgrave and Dejaeghere, 2017), with limited time or inclination to interpret and act upon financial and performance data (Moynihan, 2016; Caruana and Farrugia, 2018). Geddes (2020) reports that British MPs encounter a 'blizzard' of information and White and Hollingsworth (1999) note that, despite its intensive programme, the PAC only has time to examine about 70 per cent of the NAO's output. The Liaison Committee (2019, p. 118) suggests that 'quality' and 'timeliness' of implementation should be prioritised by committees seeking to gauge their influence on government. This provides a useful starting point but does not significantly narrow or specify the field of attention. Does 'quality' relate to government complying with the Committee's prescriptions 'to the letter', irrespective of the result? Or is it about solving the problem that

motivated the Committee's intervention, irrespective of the means taken?[5] Moreover, does timeliness refer to the pace with which solutions are implemented, how quickly positive results emerge, or the speed with which parliament is updated (thus enabling further scrutiny)? The PAC Chair, for example, recently complained that delayed reporting of regression in implementation timeframes undermines, in the words of the Treasury Officer, 'trust between the Committee and government departments as well as hindering transparency, scrutiny and accountability'.[6]

Even if information about quality and timeliness is obtained, how should performance be judged? Legislators, especially in committees with a vast remit like the PAC, may have unrealistic expectations about what is feasible in complex, resource-stretched bureaucracies (King, 1981). Equally, governments may misrepresent the difficulty of the task or urgency of its efforts (Strøm, 2000; Lupia, 2003). These problems may be partly addressed by comparing organisations undertaking similar work, in order to 'test' what is and is not possible. Thus, a second criteria for judging progress reports is whether they enable benchmarking through inter-departmental and over-time comparisons.

Finally, research into performance management regimes indicates that once a hitherto unmeasured activity is subject to monitoring and implicit or explicit rewards and sanctions, dysfunctional behaviour and 'gaming' may occur (Propper and Wilson, 2003; Bevan and Hood, 2006; Kelman and Friedman, 2009). This is succinctly expressed as Campbell's Law, whereby: 'The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor' (Campbell, 1979, p. 85). Thus, a third criteria is the extent to which the new reporting arrangements are protective against efforts by auditees to present their achievements more favourably than warranted by objective results.

4. Data preparation

To evaluate the progress reports on the basis of these three criteria, we built a dataset tracking recommendations from all PAC inquiries undertaken in the 2010 to 2012 parliamentary session across the first 14 progress reports issued up to November 2020. Some recommendations are implemented more rapidly than others, resulting in an 'unbalanced' panel consisting of 1,403 observations on our

---

[5]Other research measures the *availability* of evidence suggesting implementation has occurred, rather than the *quality* of the implementation itself (Benton and Russell, 2013).

[6]'Dear Accounting Officer' letter DAO 02/20.

main metric (implementation status).[7] (For the final set of analyses (Figure 5), we increase the sample size with additional recommendations from the 2012 to 2013 session.)

Our two main data sources are the 16 'Treasury Minutes' published between December 2010 and November 2012, which provide the government's formal response to the inquiries of 2010–12; and the 14 progress reports issued between January 2012 and November 2020. These are highly structured documents, with consistent formatting, sectioning, subtitling and indexing (see Online Appendix B), enabling precise data entry through cross-matching across multiple editions. Occasional inconsistencies were resolved by consulting the original inquiry reports or committee staff.

To capture this information, we established a dataframe in which: rows represent individual PAC recommendations; rows are grouped by originating inquiry; and columns carry inquiry- and recommendation-level observations. (Thus, the data are hierarchically structured, with some attributes shared by all recommendations from the same inquiry.) Working to a detailed protocol, a team of research assistants populated this dataframe by:

- First, obtaining details of all inquiry reports for the session (titles, publication dates, responsible departments);
- Second, inputting the PAC's recommendations underneath each report, as well as the conclusion leading to the recommendation;
- Third, capturing the government's initial response to each recommendation (date, acceptance or not, timeframe for implementation [if any]); and
- Fourth, searching all subsequent progress report for entries on each outstanding recommendation, capturing implementation status and accompanying description, key dates of achievements and/or future timeline. (Date formats were subsequently standardised according to the protocol in Online Appendix C.)

With the dataframe complete, each PAC recommendation was read and coded for the 'target organisation' required to take action; and, following Benton and Russell (2013), for the type of action recommended. Initially, we replicated exactly Benton and Russell's typology, but, after a coding trial, this was modified to reflect the PAC's unique remit, arriving at the following six recommendation categories:

- issue **guidance** and **control**;
- **clarify** the government's position and/or **disclose** information;
- undertake **analysis**, **research** and/or **data** collection;

---

[7]Total observations are the count of every unique update for the 615 recommendations published in initial government responses or any subsequent progress report.

- improve **contract management** of government suppliers;
- improve **internal management** procedures;
- **other**.

We also identified recommendations that are simply statements or conclusions, rather than actionable suggestions (these are removed from some analyses below).

Using the protocol summarised in Online Appendix C, one co-author coded the entire dataset according to this scheme, and marked each code as either a 'centre' or 'edge' case, the latter (30 per cent) indicating recommendations that potentially met criteria for two or more codes. The second co-author then blind double-coded a randomly selected 15 per cent of centre cases, as well as 100 per cent of edge cases. For centre cases, an agreement of 90 per cent was reached without discussion, with a Cohen's Kappa value of 0.881. Thus, the first co-author's centre codes were retained unmodified. For edge cases, the initial agreement was 70 per cent, with a Kappa of 0.634. Reconciled edge codes were then agreed through discussion for all disputed cases.[8]

5. Results

Figure 1 depicts the three stages of PAC oversight (left to right) and the main recommendation attributes (top to bottom), with case numbers in parentheses. Given the imbalanced nature of the panel, the variety of actions recommended by the Committee, and departmental idiosyncrasies, case numbers decline from left to right, creating a kind of 'survivorship' bias. For example, of 524 fully or partly accepted recommendations, just 240 have an explicit timeline for implementation, of which just 45 were revised one or more times. The tree diagram thus helps contextualise the varying sample sizes below.

Stage 1. Inquiry

The PAC issued 88 reports during 2010–2012, containing between 3 and 11 recommendations each (mean: 7), amounting to 615 in total. Seventy-eight of these reports (89 per cent) drew upon one or more prior studies by the NAO.

Figure 2a distributes the 615 recommendations across the 20 responsible organisations receiving a report in 2010–12 (see Online Appendix A for definitions of acronyms). As this reveals, the PAC's attention is distinctly skewed.

---

[8]By convention, Cohen's Kappa coefficients of >0.81 are typically interpreted as 'near perfect' agreement, while coefficients of 0.61–0.80 indicate 'substantial' agreement (Landis and Koch, 1977; Altman, 1999).
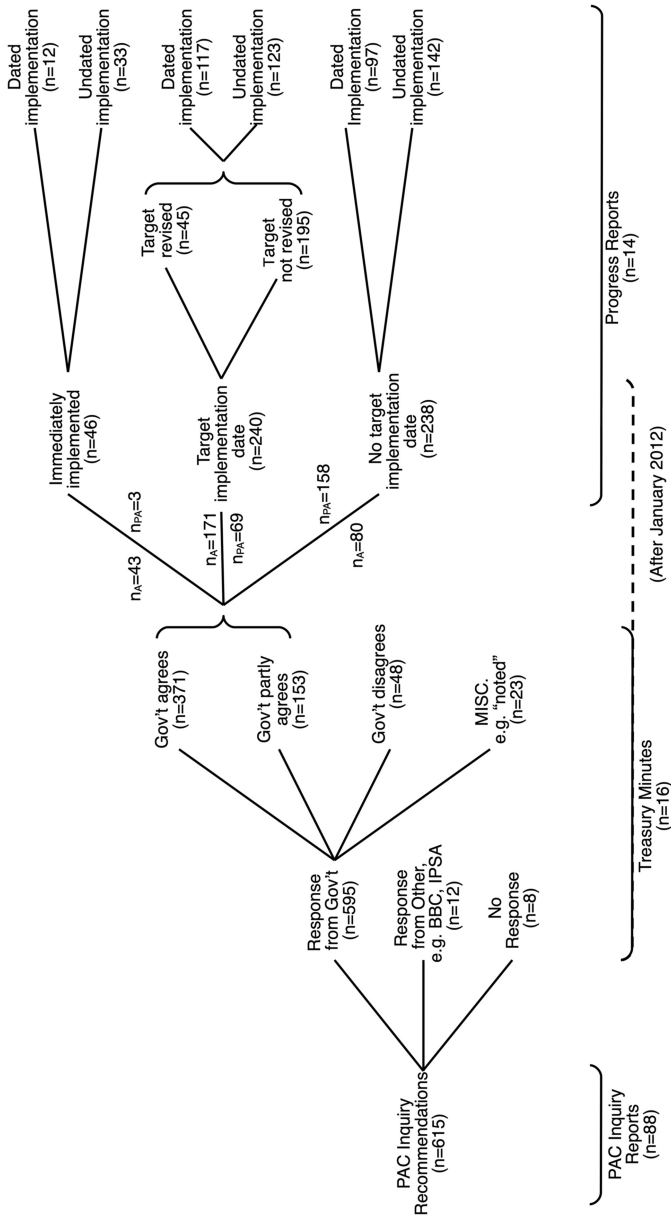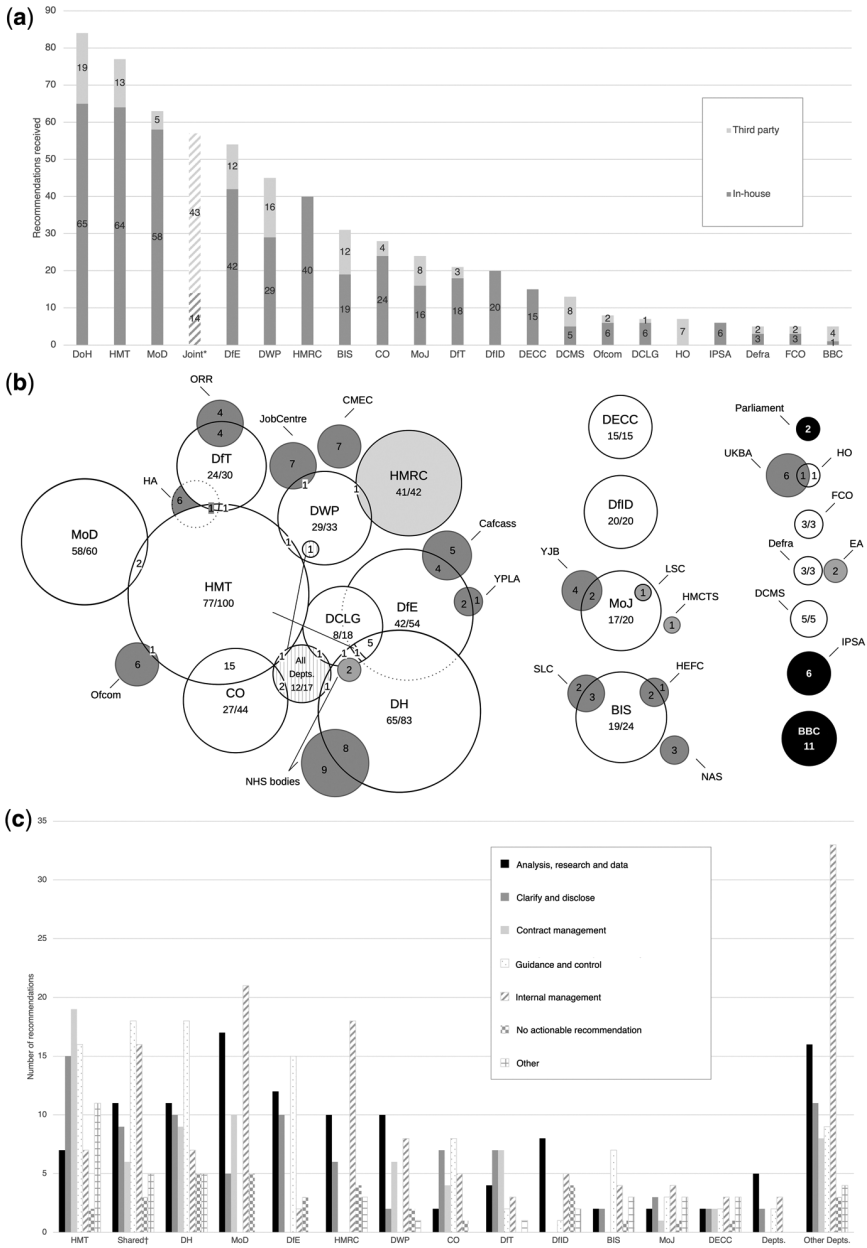
Figure 1. Acceptance of PAC recommendations, 2010–12, and implementation of accepted recommendations, 2012–20.

Figure 2. (a) Recommendations by responsible department and in-house versus third-party delivery. (*Joint includes PAC reports directed at two or more responsible departments.) (b) Recommendations by target organisation. (c) Recommendation types for most reported-upon departments. (†Shared includes PAC recommendations directed at two or more target organizations.)
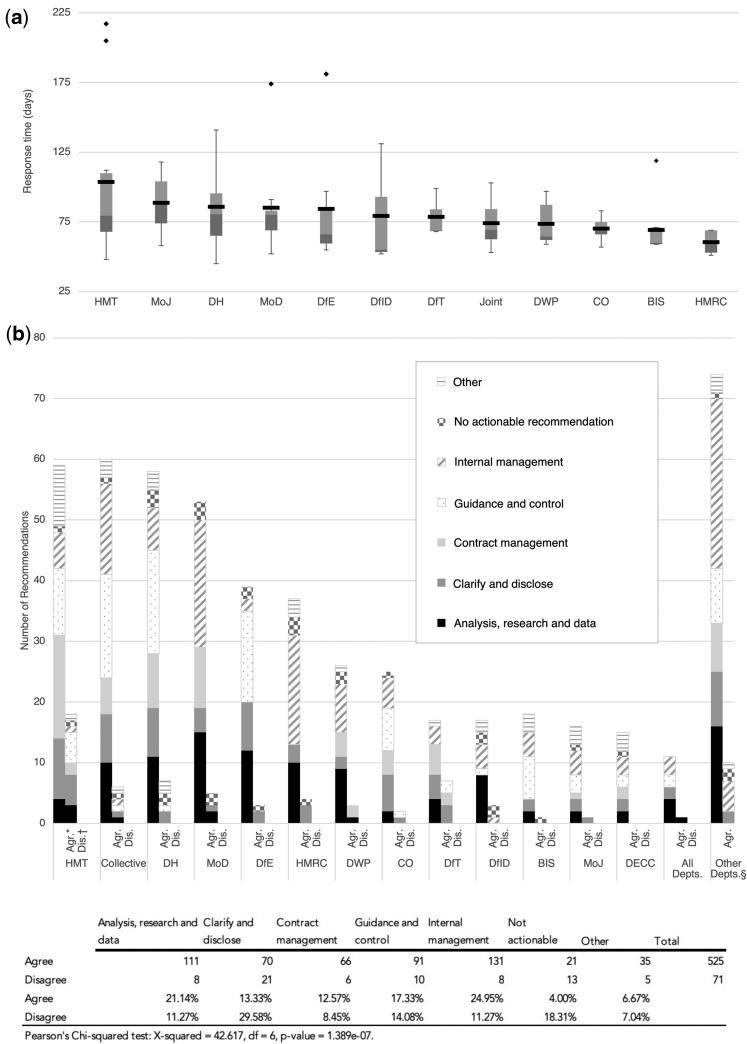
Excluding 'joint' reports to multiple departments (hashed shading), the Department of Health and HM Treasury account for nearly 30 per cent of recommendations, possibly reflecting the extensive and contentious re-organisation of the National Health Service at this time (Elston and Bevan, 2020), as well as the onset of the Treasury-managed programme of austerity which saw Whitehall undergo one of the longest periods of public spending restraint in a century (Hood and Himaz, 2017). Defence, Education, Work and Pensions and Revenue and Customs make up the next 30 per cent, with the remaining 40 per cent spread across 14 other organisations.

Expansive as this span of attention is, it is actually an underestimate, obscuring the range of 'third-party' organisations receiving recommendations (light grey element of each bar), the impact of assigning 'shared' recommendations to multiple departments, and cases where departments receive a recommendation despite not having supervisory 'responsibility' for the policy area. Consequently, the Venn diagrams in Figure 2b further breakdown the PAC's attention according to the specific organisations named in each individual recommendation (hereafter the 'target organisation'). Circle diameters are proportional to the volume of recommendations received; overlaps indicate shared recommendations directed to multiple organisations; and shading depicts 'arm's-length bodies' (non-ministerial departments, executive agencies, etc.).

This more granular analysis nearly doubles the number of organisations subject to PAC interventions, reflecting the still-fragmented organisational structure of Whitehall at this time after the decentralisation of the 1980s–90s (Elston, 2013). The figure also reveals the degree of joint working required to fulfil the Committee's suggestions. Unsurprisingly, HM Treasury forms the most connections, including 15 shared actions with the Cabinet Office (with which it jointly housed the Efficiency and Reform Group). Other departments with extensive service delivery chains, such as Health, Education and Local Government, also make multiple overlaps.

Lastly, Figure 2c compares recommendation types across departments. Of the categories drawn from the Benton and Russell (2013) coding scheme, analysis ranks highest (20 per cent, compared with 23 per cent in their study), with guidance and control second (17 per cent, compared with just 4 per cent). But, perhaps reinforcing the differing 'policy' and 'implementation' remits of departmental committees and the PAC, our most common code is internal management (23 per cent), which was novel for our study.

The distribution across departments partly conforms to expectations: improved guidance and control is prevalent among the core executive (Treasury, Cabinet Office) and departments regulating large delivery networks (Health, Education); while analysis is common among social policy centres (Education, Work and Pensions, International Development) among others (like Defence).

	Analysis, research and data	Clarify and disclose	Contract management	Guidance and control	Internal management	Not actionable	Other	Total
Agree	111	70	66	91	131	21	35	525
Disagree	8	21	6	10	8	13	5	71
Agree	21.14%	13.33%	12.57%	17.33%	24.95%	4.00%	6.67%	
Disagree	11.27%	29.58%	8.45%	14.08%	11.27%	18.31%	7.04%	

Pearson's Chi-squared test: X-squared = 42.617, df = 6, p-value = 1.389e-07.

Figure 3. (a) Response times for most reported-upon departments, plus joint reports. (Outliers are calculated as 1.5 times the inter-quartile range (IQR) above the third quantile or below the first quantile.) (b) Acceptance rates by department and recommendation type. (*Agree (Agr.) includes partly agree; [†]Disagree (Dis.) includes 'noted' etc.; [§]Other Depts. include small departments but excludes independent bodies like BBC and IPSA.)

Stage 2. Response

As shown in Figure 3a, average response times for departments to accept or reject PAC recommendations vary between 60 and 104 days. Outliers share no common attribute; and, with an acceptance rate similar to the whole sample, nor does it seem that government agonised over these delayed responses. Still, this variance

means that subsequent calculations of pace of implementation will be affected by whether counting starts from the issuing of the recommendation or, as called for in the PAC debate (see epigraph), its subsequent acceptance. (Hereafter, we adopt the latter approach.)

Of the 615 recommendations in total, 595 received a response from the government, with the remainder replied to by independent bodies, like the BBC (see Figure 1, previously). Government fully or partly agreed with 371 and 153, respectively, and explicitly disagreed with 48. The remaining 23 received non-committal responses, mainly because the PAC's comments were unactionable. This results in an acceptance rate (full or part) of 88 per cent, compared with 40 per cent for the seven departmental committees studied in Benton and Russell (2013).

Figure 3b graphs acceptance rates by recommendation type and department. To test whether some types of recommendation are more likely to be accepted, we performed Pearson's Chi-square test with the null hypothesis H_0 that the dummy agree/disagree variable is independent from recommendation type. If H_0 holds, the agree/disagree distribution across recommendation types should be proportionate to the share of recommendations in each category. In fact, the observed distribution is not the same as our expected distribution (significant at $P = 0.000$ level). Therefore, we reject H_0 and accept that there is a relationship between recommendation type and acceptance. Conclusions with no actionable element are disproportionately more likely to be rejected, as are requests to release or clarify existing information. Conversely, recommendations for more analysis or improved internal management or contract management are less likely to be rejected. The willingness to commit to research, for example, may reflect the legitimacy attached to notions of 'evidence-based' policymaking (Boswell, 2009), or be because further research could still eventually result in no change in government position and is thus an 'easy' concession to make.

Beyond recommendation type, it was not possible to robustly measure the magnitude of the actions requested by the PAC, although prior research indicates that government is more likely to accept recommendations by departmental committees when smaller changes are required (Benton and Russell, 2013).

Stage 3. Progress

Turning to the novel 'progress reporting' stage of oversight, Figure 4a presents the most straightforward analysis possible from the new data: the final status of each recommendation after 14 rounds of reporting. The overwhelming majority are self-reported as implemented by this point. Typically, this is conveyed in the progress reports (85 per cent), but sometimes completion is notified in the government's initial response (9 per cent). In 14 cases (2.7 per cent), government
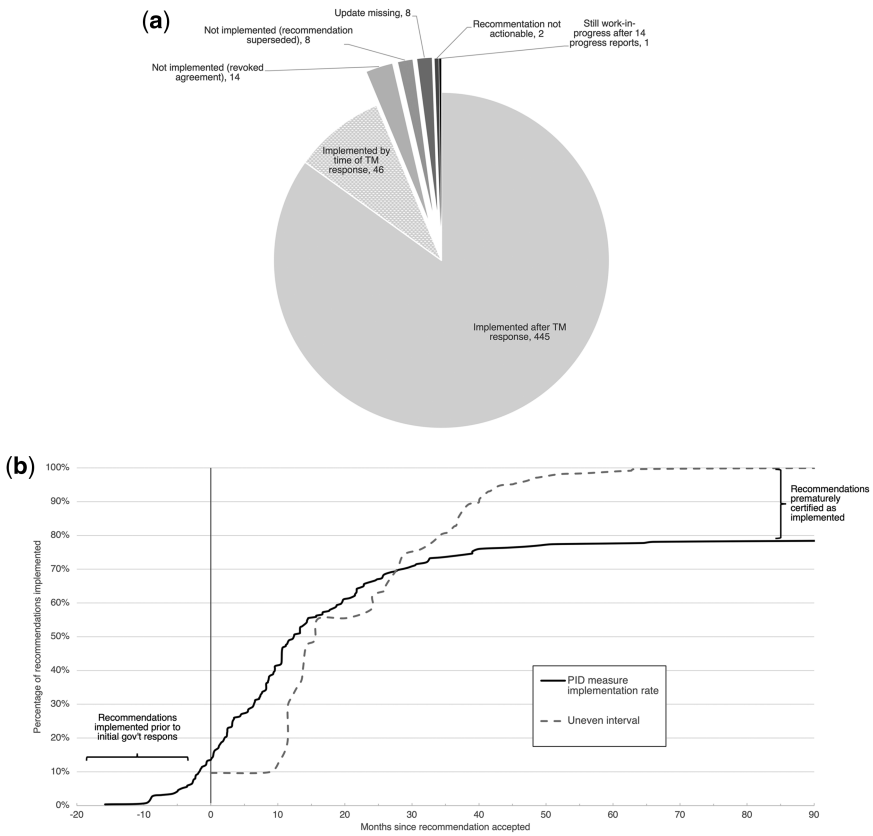
Figure 4. (**a**) Status of 2010–12 recommendations by November 2020. (**b**) Implementation rate comparing approximate and precise measures of time to completion. (**c**) Implementation rate by recommendation type. (**d**) Target versus actual implementation dates (continued on next page).

subsequently revoked its initial acceptance and aborted implementation of a recommendation, and a further eight suggested actions (1.5 per cent) were superseded by new reports or policy changes. Only eight times did the government fail to provide a scheduled update, and only one recommendation (concerning debt collection by the Department for Education) is still 'work-in-progress' after 14 reports.

As for the pace of implementation, there are three possible ways to measure this. First is to count the number of progress reports published between accepting and completing each recommendation.[9] More than 60 per cent of actions are

---

[9]Given the intermittent coverage of PAC inquiries in the early iterations of the progress reports, we count the number of progress reports between acceptance and implementation that actually *mention* the original PAC inquiry and thus provide an opportunity for government to update the Committee.

(c)

(d)
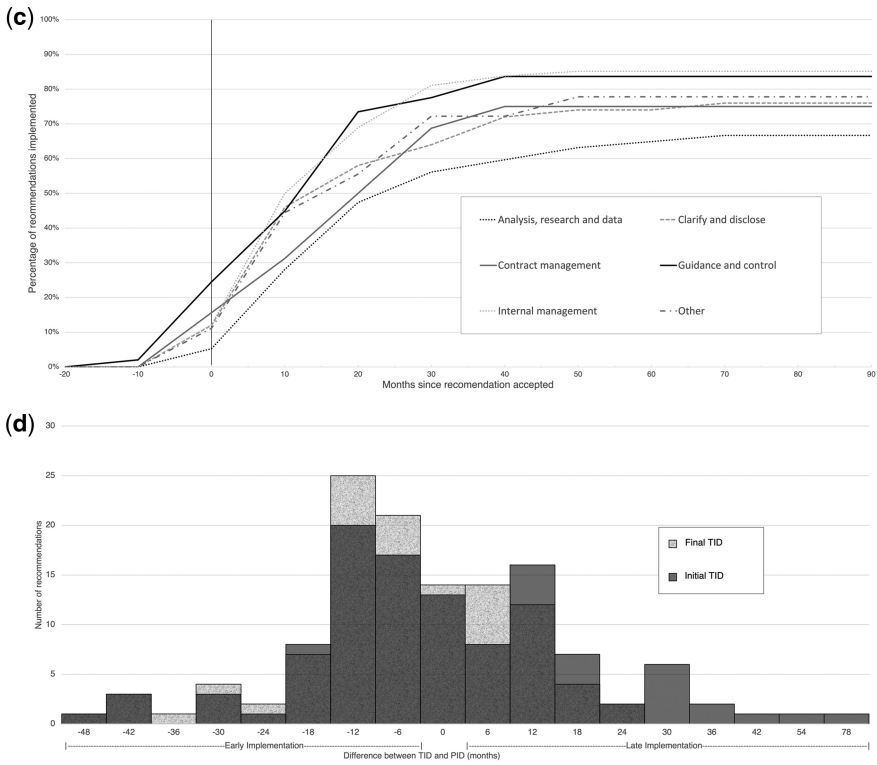
Figure 4. continued

completed by the first progress report, rising to 95 per cent by the third. However, though straightforward to undertake and understand, this method is flawed, mainly because of the uneven interval between each publication (varying from 3 to 18 months). An action completed in its first progress report but in month 17 of an 18-month reporting window would be ranked higher than one completed after two progress reports issued in the space of 12 months.

Second is to measure time to reported completion, taking into account the unique combination of reporting intervals for recommendations accepted at different times during 2010–12. This is plotted with the dashed grey line in Figure 4b. Now, 60 per cent completion is reached at the 25-month mark, while 95 per cent is achieved at about 45 months. Still, this approach is problematic, however, since actions achieved at different ends of the same reporting window (e.g. months 1 and 11 of a 12-month period) are unfairly equated. (This would be less problematic if reporting windows were short and equally sized.)

Overcoming these problems requires manual recording of any precise implementation date provided in the narrative description accompanying the final entry for each recommendation. Nearly half of cases contain no such date, reducing the sample to $n = 226$. The continuous black line in Figure 4b depicts the result. For this subsample, 60 per cent completion is achieved slightly earlier (at 20 months), but the 95 per cent mark is never reached. This is because, despite the recommendation being classified as implemented and appearing in no future progress reports, in 62 cases (21.5 per cent) the completion dates provided are *prospective*, citing the expected date that outstanding work will be completed.

For example, in November 2011, the PAC recommended that the Department of Health clarify its rules regarding the use of personal budgets by users of social care, and develop a complaints process. The department revised its intended timeline three times, finally arriving at a target of April 2016. As this date approached, however, the promised new appeal system was dropped, and completion of the action was declared purely on the basis of the requested clarification, with a promise of future work on appeals by April 2020, more than 4 years hence.

Not all cases of premature self-certification are as egregious as this, and some may be (rationalised as) an innocent shortcut intended to limit excessive reporting. But this practice also releases departments from the obligation to provide a final update about what is possibly the hardest outstanding elements of completion, and thus may facilitate gaming through either performance-inflation or scrutiny-avoidance.

Using the most robust but labour-intensive method of measuring pace, developed in Figure 4b, Figure 4c compares implementation trajectory across recommendation types. The internal management and guidance and control types achieve both the fullest and most rapid implementation, possibly reflecting the low dependence of such actions on third parties. Recommendations regarding contract management and the clarification of policy positions or disclosure of information occupy an intermediate position. Here, interdependencies with contractors (who must agree to contractual changes) or ministers or special advisers (regarding policy transparency) might delay or inhibit progress. Lastly, analysis, research and data achieves the slowest trajectory and lowest completion rate, reflecting the long lead time for research and, again, the use of prospective completion dates.

As well as the pace of implementation, Figure 4b and c also speaks to debates about the unique impact of oversight given that many of the recommendations may already have been under consideration within the audited organisations (Hindmoor *et al.*, 2009). Some 8 per cent of recommendations are implemented even before government provides its initial accept/reject response to the Committee. Moreover, the earliest recorded implementation date—concerning

(a)

(b)

(c)
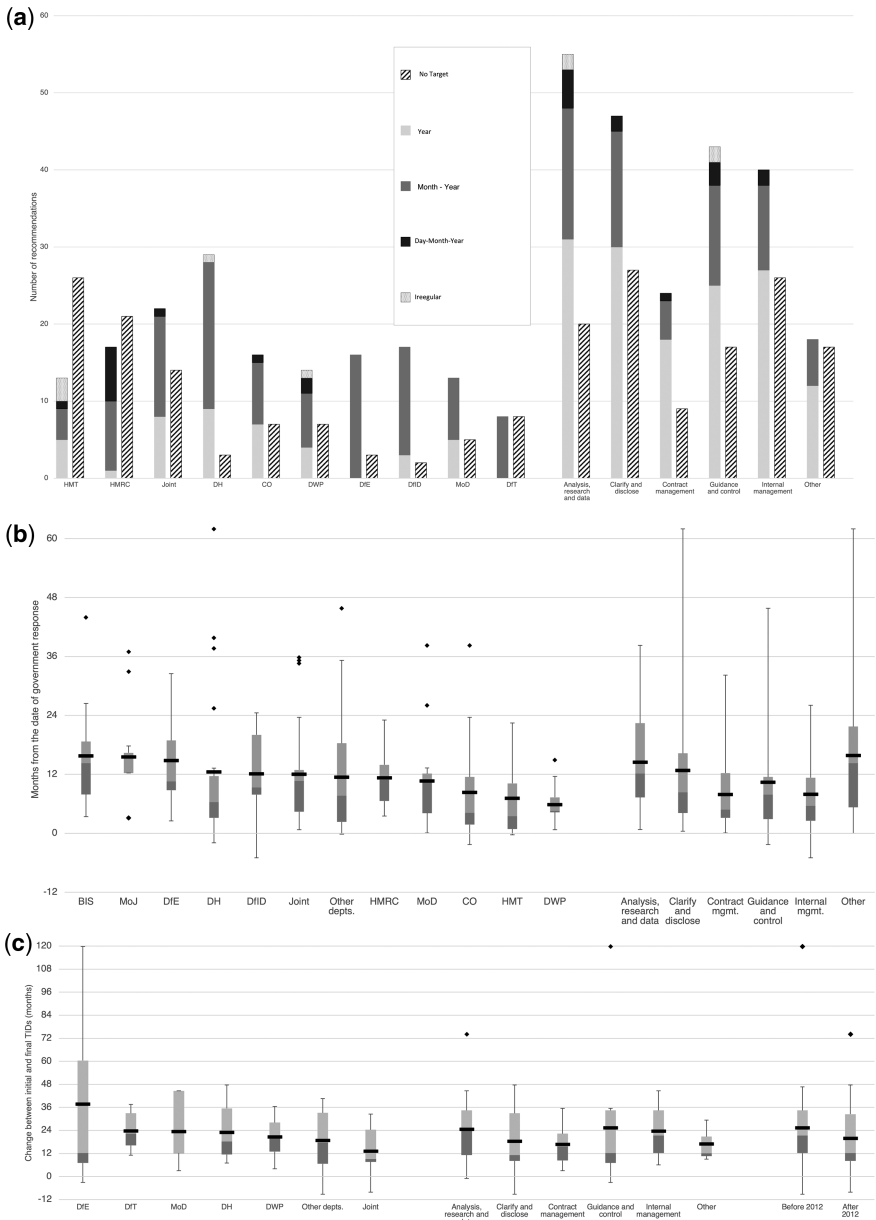
Figure 5. (a) Precision of target dates by major organizations and recommendation type, 2010–2013. (b) Duration of initial target dates by major organizations and recommendation type, 2010–2013. (For recommendation types, outliers are suppressed and whiskers indicate true range to aid visualization.) (c) Revisions to target dates by major organisations, recommendation type and date of response, 2010–2013.

spending controls used by the core executive during austerity—occurred some 16 months *prior* to the government's response to the Committee and 13 months prior to its inquiry. Cases such as this rule out the direct influence of the Committee on government policy, but not the indirect 'deterrence' effects mentioned previously.

Aside from the pace of implementation, Figure 4d compares the target implementation dates ('TIDs') provided by government with manually-coded dates of completion. Since this analysis relies on cases possessing *both* an explicit timeline and specified implementation date, the sample size is further reduced ($n = 110$). Moreover, because departments have the flexibility to revise TIDs during implementation (a cause of complaint by the PAC chair), the figure contains two overlapping histograms: one comparing performance against initial TIDs (dark grey) and the other using final TIDs (light grey, textured).[10] The data approach a bell-shape distribution around a mode completion time of 6–12 months ahead of schedule, with 44.5 per cent of the sample completed within 6 months (early or late) of the final timetable and 78.2 per cent within 12 months. As might be expected, the effect of revising the timetables during execution is to improve government's scorecard considerably, with 11 extra recommendations completed within 6 months of the (modified) target and no action more than 24 late.

Our final set of analyses explores target setting in more depth, using additional recommendations from 2012 to 2013 to increase case numbers.[11] The left side of Figure 5a records the precision with which departments set their timetable and the frequencies with which TIDs are omitted. Several inter-departmental differences emerge. Both the Treasury and Tax departments often avoid committing to a target altogether (hashed bar); whereas this is rare or very rare for other ministries, especially Health, Education and Development. Moreover, about half of the Treasury's targets contain minimal precision (slightly less so for Cabinet Office and Defence), providing officials with more 'wiggle room' before actions fall late. Conversely, the greatest precision (Month–Year; Day–Month–Year) occurs with the two departments whose operations are most obviously calendar-driven (financial year, for HMRC; school year, for Department for Education). There is no obvious association between timeline precision and recommendation types.

Figure 5b summarises the length of initial timelines set by each department and for each recommendation type. Means range between 6 and 16 months,

---

[10]When no revision is made to the TID, the same date is used for the initial and final TID analyses.

[11]Since 2012, departments have assigned TIDs at the point the recommendation is accepted. This is then either reiterated or revised during the subsequent progress reports. Recommendations accepted prior to December 2011 did not receive a timeline at the point of acceptance and only have a TID if still categorised as work-in-progress during their first appearance in a progress report after January 2012, reducing the number of analysable cases considerably (see Figure 1).

somewhat skewed by outliers, especially in Health. Despite being reluctant to set (precise) targets, HM Treasury achieves the lowest-but-one average for its timelines (just 7 months), compared with the much more target-prone Education and Health (15 and 13 months, respectively). This could reflect this distinct nature of the finance ministry's recommendations; but it may also indicate 'cream skimming', whereby commitment to a timeline is only given for relatively straightforward actions with a low risk of failure. As for recommendation type, the average committed timelines are similar, with 'clarify and disclose', 'guidance and control' and 'other' recommendations having larger variation.

Lastly, for those cases ($n = 63$, including the 2012–2013 data) in which departments revised their TIDs once or more, Figure 5c compares the regress between initial and final timelines. The average total effect of the revisions ranges between 1 and 3 years. As the comparison between the 'before' and 'after 2012' groups shows, departments were initially less skilled at estimating timeframes and thus required longer extensions for recommendations accepted prior to the reforms, compared with subsequently.

6. Discussion

6.1 Performance against criteria

Based on the analyses presented above, how well do the progress reports meet the aforementioned criteria of: (1) informing parliament about the actual—not merely intended—speed and quality of implementation, (2) enabling over-time and inter-organisational comparisons, and (3) suppressing opportunities for gaming?

The data indicate that the majority of accepted recommendations are self-declared to be complete; although only about 50 per cent are within Margaret Hodge's 12-month deadline (see epigraph). In 85 per cent of cases, notification was issued in the progress reports rather than the initial response to the Committee, implying that the new arrangements are improving post-inquiry transparency (though we lack a pre-reform comparator). But the extent to which the reports merely consolidate and curate disparate information already in the public domain, as opposed to releasing genuinely new data, is unclear. Updates on each recommendation are typically retrospective in nature, identifying action that has been taken since the last report was issued. But there are significant exceptions to this, as indicated by the 62 cases in which actions were closed despite departments issuing prospective completion dates, sometimes several years ahead.

The arrangements score less highly at enabling the comparisons necessary to benchmark departmental performance or form broader conclusions about

responsiveness. Indeed, the reports serve mainly a 'look up' or 'reference' function, enabling departments to set out the steps taken in response to PAC inquires, and streamlining (to a degree) the process by which interested parties can form an opinion about action taken in respect of any particular recommendation. The non-digital format (in contrast with the searchable/filterable recommendation trackers of both the NAO and the Government Accountability Office in the USA) means that it is difficult to 'keep score' of departments or recommendation types without an extensive database building exercise (something that both PAC staff and the NAO have attempted). And even when such investment is made on a moderate scale (in our case, more than 600 recommendations tracked over 8 years), the irregular publication schedule and inter-departmental inconsistencies in reporting style impede comparative analysis.

Finally, no system of performance monitoring is entirely immune to gaming, and nor does the occurrence of strategic behaviour necessarily invalidate the measurement exercise (Kelman and Friedman, 2009). Nevertheless, the current design of the progress reports does create an unnecessary degree of vulnerability to gaming. As we have seen, failure to provide a timeline for implementation, and/or opting for lengthier and less precise targets, allows some departments to absolve themselves from demanding schedules. Similarly, premature classification of actions as completed stymies the scrutiny process and misleads parliament as to the pace of implementation and (potentially) the ultimate outcome. And the possibility of delaying publication of an update until a favourable account can be made, either by omitting a scheduled update from a progress report or postponing publication of the entire progress report, further risks the credibility of the reporting system.

6.2 Improvements and extensions

It is premature to draw conclusions about whether the progress reporting should continue, given the research that needs to be undertaken to understand the use and impact of this data (discussed in the 'Conclusion' section, below). But there are clearly refinements that should be made in the interim to strengthen usability and usefulness, and reduce the risk of gaming. Firstly, as both an aid to interpretation by the PAC and as a deterrent against misreporting, the NAO should audit a sample of information contained in the progress reports. Consistent reporting intervals, a transparent and immovable publication schedule, and a requirement to, whenever meaningful,[12] provide precise completion dates would also help combat many of the problems noted above. Greater insistence on departments

---

[12]Precise dates may be omitted because: they are inappropriate for some types of recommendation (e.g. achieving 'culture change' among staff); they are alien to the drafting style of the department; or to obscure poor performance.

setting and justifying timelines, and regulation of when and how these can be re-vised, is necessary. And the Treasury should develop, publish and enforce explicit guidelines on the circumstances in which accounting officers can self-report actions as 'implemented' or 'completed', and on ways in which departments should indicate what elements of the PAC's recommendation are *not* being taken forward or have been dropped during the implementation process, since this information is easily lost within the overall narrative. For example, early in 2012, government was asked to release the information it relies upon to evaluate local flood authorities' risk management strategies. Despite being a relatively straight-forward 'disclosure' request, a vague and distant target date three years hence was provided, by which time the department promised to have completed such an evaluation. However, the action was closed early with no such review having been undertaken, and with the department simply pointing to material already in the public domain as fulfilling its obligations. Such examples illustrate the flexibility with which departments report actions as completed—even for cases of signifi-cant under-delivery.

Looking beyond the PAC, the Liaison Committee's (2019, p. 23) recent recom-mendation that 'every government department [produce] an annual memorandum to the relevant select committee which sets out progress on implementing commit-tee recommendation' raises the question of whether the progress-reporting system developed by the Treasury should be extended to other committees. Differences in context would need to be considered; in particular, the lower volume of government-accepted recommendations on which to report. In addition, central co-ordination within government and/or parliament would be desirable to ensure com-parability of data across policy sectors and appropriate control over the frequency and form of departmental responses, pre-empting the problems highlighted above. And the question of overload needs to be addressed. As Brazier and Fox (2011, pp. 360–361) note, further expansion of select committee responsibilities without sub-stantially increasing resources may not be feasible or desirable. An intermediate step might be to seek greater use of the data already being produced by the Treasury Officer of Accounts by asking departmental committees to scrutinise their own charges' follow-up to the PAC's work. While this would require inter-committee collaboration, the recent growth in multi-committee inquiries (Geddes, 2020) and the potential for coordination through the Scrutiny Unit suggests that this is not an unrealistic prospect.

6.3 Study limitations

Our study contains several limitations to be addressed in future research. Firstly, our data relate to an extraordinary period in British politics, conditioned by the novelty of coalition government, the scale of national austerity, recent select

committee reform following the parliamentary expenses scandal, an unusually high-profile PAC Chair and a number of significant departures from established PAC practice, including a 'ground-breaking' (Hodge, 2016, p. 48) campaign against tax evasion.[13] Our findings are not necessarily generalisable to other periods, therefore. Secondly, we lacked the resources to develop a robust method for coding the size of the changes sought by each recommendation, to verify claims made by departments through other sources, or to assess the quality of implementation based on standardised criteria. Incorporation of recently published data from the NAO, which has partially verified actions taken in respect of its recommendations since 2019, is a promising option. And thirdly, while our three evaluation criteria cover a broad spectrum of issues, they are not comprehensive. Our analysis has not considered whether and how information provided in the reports is actually used, by whom and for what purpose; and nor, crucially, has it explored the impact on auditees.

7. Conclusion and research agenda

Many have cautioned that insufficient attention to following-up prior investigations will unnecessarily curtail the influence of select committees on government. 'The impact of committee reports', wrote the Hansard Society (2001), 'will be determined by the assiduity with which their recommendations are monitored'. Much of the literature on public audit concurs, with the Audit Forum even suggesting that 'reporting without any follow-ups is similar to doing no audit at all' (cited in Kayrak, 2008). Nonetheless, few examples of 'recommendation tracking' have been forthcoming internationally, and there has been limited debate about how such work should best be undertaken. Moreover, existing comparative literature on public accounts committees in Westminster-style parliamentary systems relies heavily on small, perception surveys of PAC officials (Wehner, 2003), whose accounts of the implementation process are rarely supported by recommendation- or inquiry-level data (Pelizzo *et al.*, 2006).

We have begun to address this gap by exploring eight years of data on the implementation of Public Accounts Committee recommendations made available following a little-known reform in 2012. Our main finding is that the progress reports are presently 'data rich but information poor', given the difficulties of drawing inferences about trends and patterns. Furthermore, we have presented initial evidence that: departments with fewer target dates may be more ambitious in setting those timelines; timeline revisions serve to limit project overruns to a maximum of two years; recommendation type may be associated with the pace of implementation; and (in)actions

---

[13]During 2010–15, 14 out of 16 PAC inquiries focusing on tax issues centred around topics including tax fraud, tax credit benefit abuse, tax evasion and tax collection enforcement. Conversely, during 2005–10, only 1 out of 14 reports examined these topics.

can be reported to the Committee up to 19 months after they occurred (or did not occur). Still, our research was intentionally exploratory, and as such raises more questions than it answers, which we outline now by way of conclusion.

7.1 Future research

Our first set of suggestions is for research into the origins and impact of the 2012 reforms. How did government respond to the prospect of greater transparency following the PAC motion in December 2010, and how was the reporting format agreed over the subsequent 13 months? Was the Treasury enthusiastic, viewing the process as a new tool to bear down on departmental expenditure during the 'age of austerity'?[14] And, commensurately, were departments resistant to the change? Has the reform measurably altered departmental behaviour in a desirable or undesirable way? For instance, has the willingness to accept recommendations in the first place, or the attributes of the recommendations that are accepted, altered? Do accounting officers discuss more fully the PAC's work and their departmental response within departmental boards and audit committees? And, outside of government, how frequently, in what context, and for what purposes are the progress reports used by, say, MPs, their staff, the media or interest groups?

Second is to explore factors affecting the pace and quality of implementation, and, therein, conditions in which the executive is more or less responsive to the wishes of the legislature. For instance, to what extent does the quality and authority of committee recommendations, and the audits that underpin them, determine their eventual implementation? (Hindmoor *et al.*, 2009) Are there organisation-level factors, such as propensity for organisational learning (Stark, 2019), decision-making structures and political (in)dependence (Elston, 2013), or internal resources and capabilities (Bolton *et al.*, 2016; Drolc and Keiser, 2021), that influence pace and quality? What is the effect of assigning 'shared' recommendations to multiple organisations, or of suggesting changes that require coordinated action across networks of decentralised bodies? Is there a trade-off between speed and quality of implementation, as often arises in bureaucracies (Hood, 1976)? And how do disruptive events, such as Brexit or the Covid-19 pandemic, affect implementation? Finally, what effect, if any, does progress reporting have on the executive's ability to address recurrent themes in the PAC and NAO inquiries, such as government's inadequate contract management capability?

Third is to explore ways to integrate the PAC data with that relating to other committees (e.g. Benton and Russell, 2013), as well as with the NAO's recent

---

[14]We are grateful to the anonymous reviewer for this suggestion.

recommendation tracker.[15] What empirical evidence is there for the claimed superior influence of the PAC vis-à-vis other committees, and is this dynamic/evolving? It may also be possible to compare the effects of the PAC and NAO working 'jointly and severally' by examining the implementation of recommendations issued by both bodies and one or the other; and, in so doing, gain new insights into the combination of political and administrative factors that explain executive responsiveness to oversight.

Supplementary data

Supplementary data are available at *Parliamentary Affairs* online.

Acknowledgements

Conflict of interest

The authors have no conflicts of interest to report.

References

Aldons, M. (2000) 'Rating the Effectiveness of Committee Reports: Some Examples', *Australian Parliamentary Review*, **16**, 52–99.

Altman, D. (1999) *Practical Statistics for Medical Research*. Boca Raton, FL, Chapman & Hall/CRC.

Anand, P. (1988) 'Monitoring and Auditing Value for Money in the UK: The Scope for Quantitative Analysis', *Financial Accountability and Management*, **4**, 253–270.

Aylett, P. (2019) 'Reform and Consolidation: A New Perspective on Commons Select Committees 1960–1980', *Parliamentary Affairs*, **72**, 742–760.

Benton, M. and Russell, M. (2013) 'Assessing the Impact of Parliamentary Oversight Committees: The Select Committees in the British House of Commons', *Parliamentary Affairs*, **66**, 772–797.

Bevan, G. and Hood, C. (2006) 'What's Measured Is What Matters: Targets and Gaming in the English Public Healthcare System', *Public Administration*, **84**, 517–538.

---

[15]https://www.nao.org.uk/nao-recommendations-tracker/.

Bolton, A., Potter, R. and Thrower, S. (2016) 'Organizational Capacity, Regulatory Review, and the Limits of Political Control', *Journal of Law, Economics, and Organization*, **32**, 242–271.

Boswell, C. (2009) *The Political Uses of Expert Knowledge*, Cambridge, Cambridge University Press.

Bovens, M. A. R. K., Schillemans, T. and Hart, P. (2008) 'Does Public Accountability Work?', *Public Administration*, **86**, 225–242.

Brazier, A. and Fox, R. (2011) 'Reviewing Select Committee Tasks and Modes of Operation', *Parliamentary Affairs*, **64**, 354–369.

Bringselius, L. (2014) 'The Dissemination of Results from Supreme Audit Institutions: Independent Partners with the Media?', *Financial Accountability & Management*, **30**, 75–94.

Campbell, D. (1979) 'Assessing the Impact of Planned Social Change', *Evaluation and Program Planning*, **2**, 67–90.

Caruana, J. and Farrugia, B. (2018) 'The Use and Non-use of the Government Financial Report by Maltese Members of Parliament', *Accounting, Auditing & Accountability Journal*, **31**, 1124–1144.

Chapman, L. (1978) *Your Disobedient Servant*. London: Chatto and Windus.

Constitution Committee. (2013) *The Pre-emption of Parliament (HL 165, Session 2012–13)*. London: The Stationery Office.

Dewar, D. and Funnell, W. (2016) *A History of British National Audit*. Oxford: Oxford University Press.

Drewry, G. (ed.) (1985) *The New Select Committees*. Oxford: Clarendon Press.

Drolc, C. and Keiser, L. (2021) 'The Importance of Oversight and Agency Capacity in Enhancing Performance in Public Service Delivery', *Journal of Public Administration Research and Theory*, **31**, 773–789.

Dubnick, M. (2005) 'Accountability and the Promise of Performance', *Public Performance & Management Review*, **28**, 376–417.

Dunleavy, P. and Muir, D. (2013) *Parliament Bounces Back—How Select Committees Have Become a Power in the Land*. Democratic Audit. Retrieved 24 March 2021 from https://www.democraticaudit.com/2013/07/18/parliament-bounces-back-how-select-committees-have-become-a-power-in-the-land/.

Elston, T. (2013) 'Developments in UK Executive Agencies: Re-examining the "Disaggregation-Reaggregation" Thesis', *Public Policy and Administration*, **28**, 66–89.

Elston, T. and Bevan, G. (2020) 'Using Opportunity Costs to Counter "One-Shot" Bias in Policy Innovation'. In Sullivan H. and Dickinson H. (eds) *The Palgrave Handbook of the Public Servant*. Basingstoke, Palgrave Macmillan, pp. 1665–1684.

Flesher, D. and Zarzeski, M. (2002) 'The Roots of Operational (Value-for-Money) Auditing in English-Speaking Nations'. *Accounting and Business Research*, **32**, 93–104.

Flinders, M. (2012) *Defending Politics*. Oxford: Oxford University Press.

Foster, D. (2015) 'Going "Where Angels Fear to Tread": How Effective Was the Backbench Business Committee in the 2010–2012 Parliamentary Session?', *Parliamentary Affairs*, **68**, 116–134.

Gaines, B., Goodwin, M., Bates, S. and Sin, G. (2019) 'A Bouncy House? UK Select Committee Newsworthiness, 2005–18', *The Journal of Legislative Studies*, **25**, 409–433.

Geddes, M. (2018) 'Committee Hearings of the UK Parliament: Who Gives Evidence and Does This Matter?', *Parliamentary Affairs*, **71**, 283–304.

Geddes, M. (2020) *Dramas at Westminster*. Manchester: Manchester University Press.

Hansard Society. (2001) *The Challenge for Parliament*. London: Vacher Dod.

Hennessy, P. (1990) *Whitehall*. London: Fontana Press.

Hindmoor, A., Larkin, P. and Kennon, A. (2009) 'Assessing the Influence of Select Committees in the UK: The Education and Skills Committee, 1997–2005', *The Journal of Legislative Studies*, **15**, 71–89.

HM Treasury. (2015) *The Accounting Officer's Survival Guide*. London.

Hodge, M. (2016) *Called to Account*. London: Little, Brown.

Hood, C. (1976) *The Limits of Administration*. London: Wiley.

Hood, C. and Himaz, R. (2017) *A Century of Fiscal Squeeze Politics*. Oxford: Oxford University Press.

Judge, D. (1992) 'The "Effectiveness" of the Post-1979 Select Committee System: The Verdict of the 1990 Procedure Committee', *The Political Quarterly*, **63**, 91–100.

Keane, J. (2009) *The Life and Death of Democracy*. London: Simon and Schuster.

Kells, S. (2011) 'The Seven Deadly Sins of Performance Auditing: Implications for Monitoring Public Audit Institutions', *Australian Accounting Review*, **21**, 383–396.

Kelman, S. and Friedman, J. (2009) 'Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service', *Journal of Public Administration Research and Theory*, **19**, 917–946.

King, A. (1981) 'How to Strengthen Legislatures—Assuming That We Want to'. In Ornstein NJ (ed.) *The Role of the Legislature in Western Democracies*. Washington, DC, American Enterprise Institute, pp. 77–89.

Kubala, M. (2011) 'Select Committees in the House of Commons and the Media', *Parliamentary Affairs*, **64**, 694–713.

Landis, J. and Koch, G. (1977) 'The Measurement of Observer Agreement for Categorical Data', *Biometrics*, **33**, 159–174.

Liaison Committee. (2019) *The Effectiveness and Influence of the Select Committee System (HC 1860, Session 2017–19)*. London: The Stationery Office.

Lonsdale, J. (2008) 'Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit', *Evaluation*, **14**, 227–248.

Lupia, A. (2003) 'Delegation and Its Perils'. In Strøm K., Müller W. C. and Bergman T. (eds) *Delegation and Accountability in Parliamentary Democracies*, Oxford, Oxford University Press. pp. 34ff.

Matthews, F. (2020) 'Parliamentary Scrutiny of Executive Patronage: The Relationship between Institutional Norms, Reputation and Accountability', *Public Administration*, **98**, 625–642.

Mattson, I. and Strøm, K. (1995) 'Parliamentary Committees'. In Döring H. (ed.) *Parliaments and Majority Rule in Western Europe*. Frankfurt, Campus Verlag. pp. 249–307.

Midgley, H. (2019) 'The National Audit Office and the Select Committee System 1979–2019', *Parliamentary Affairs*, **72**, 779–798.

Morin, D. (2001) 'Influence of Value for Money Audit on Public Administrations: Looking beyond Appearances', *Financial Accountability & Management*, **17**, 99–117.

Morin, D. (2008) 'Auditors General's Universe Revisited: An Exploratory Study of the Influence They Exert on Public Administration through Their Value for Money Audits', *Managerial Auditing Journal*, **23**, 697–720.

Moynihan, D. (2016) 'Political Use of Performance Data', *Public Money & Management*, **36**, 479–481.

National Audit Office (NAO). (2020) *Annual Report and Accounts, 2019–20. (HC 373 Session 2019–2021)*. London.

NAO. (2021) *History of the NAO*. Retrieved 24 March 2021 from: https://www.nao.org.uk/about-us/our-work/history-of-the-nao/.

Pelizzo, R., Stapenhurst, R., Sahgal, V. and Woodley, W. (2006) 'What Makes Public Accounts Committees Work? A Comparative Analysis: What Makes Public Accounts Committees Work?', *Politics & Policy*, **34**, 774–793.

Pollitt, C., Girre, X., Lonsdale, J., Mul, R., Summa, H. and Waerness, M. (1999) *Performance or Compliance*. Oxford: Oxford University Press.

Powell, G. (2000) *Elections as Instruments of Democracy*. New Haven, CT: Yale University Press.

Power, M. (1997) *The Audit Society*. Oxford: Oxford University Press.

Procedure Committee. (2019) *Should There Be a Commons Budget Committee? (HC 1482, Session 2017–19)*. London: The Stationery Office.

Propper, C. and Wilson, D. (2003) 'The Use and Usefulness of Performance Measures in the Public Sector', *Oxford Review of Economic Policy*, **19**, 250–267.

Public Accounts Committee (PAC). (2007) *Holding Government to account: 150 Years of the Committee of Public Accounts*. London.

PAC. (2015) *The Work of the Committee of Public Accounts 2010–15 (HC 1141)*. London: The Stationery Office.

Rhodes, R. (2011) *Everyday Life in British Government*. Oxford: Oxford University Press.

Robinson, A. (1985) 'The Financial Work of the Select Committees'. In Drewry G. (ed.) *The New Select Committees*, Oxford, Clarendon Press, pp. 307–318.

Rogers, R., Walters, R., Besly, N. and Goldsmith, T. (2019) *How Parliament Works*. London: Routledge.

Russell, M. (2011) 'Never Allow a Crisis to Go to Waste': The Wright Committee Reforms to Strengthen the House of Commons', *Parliamentary Affairs*, **64**, 612–633.

Russell, M. and Gover, D. (2017) *Legislation at Westminster*. Oxford: Oxford University Press.

Staddon, A. (2015) 'The Public Accounts Committee of the House of Commons'. In Hoque Z. (ed.) *Making Governments Accountable*. Routledge.

Stark, A. (2019) 'Policy Learning and the Public Inquiry', *Policy Sciences*, **52**, 397–417.

Strøm, K. (2000) 'Delegation and Accountability in Parliamentary Democracies', *European Journal of Political Research*, **37**, 261–290.

Talbot, C. (2010) Performance in Government: The Evolving System of Performance and Evaluation, Measurement, Monitoring and Management in the United Kingdom: World Bank, ECD Working Paper Series No. 24.

Talbot, C. and Wiggan, J. (2010) 'The Public Value of the National Audit Office', *International Journal of Public Sector Management*, **23**, 54–70.

Walgrave, S. and Dejaeghere, Y. (2017) 'Surviving Information Overload: How Elite Politicians Select Information', *Governance*, **30**, 229–244.

Wehner, J. (2003) 'Principles and patterns of financial scrutiny: Public Accounts Committees in the Commonwealth', *Commonwealth & Comparative Politics*, **41**, 21–36.

White, F. and Hollingsworth, K. (1999) *Audit, Accountability and Government*. Oxford: Clarendon Press.