

A Global Mapping of Delivery Approaches

Zahra Mansoor
Dana Qarout
Kate Anderson
Celeste Carano

Liah Yecaló-Teclé
Veronika Dvorakova
Martin J. Williams

Acknowledgements

The team would like to acknowledge the support and contributions of colleagues across the Blavatnik School of Government at the University of Oxford and the Education Commission. Mansoor, Qarout, Yecalotecle, Dvorakova, and Williams are affiliated with the University of Oxford; Anderson is affiliated with the Education Commission; Carano was formerly affiliated with the Education Commission and contributed until November 2020.

We are grateful for the helpful comments and inputs from the [DeliverEd High-Level Advisory Group](#), in particular co-chairs Amel Karboul and Mercedes Miguel and special advisor to the DeliverEd Initiative Gwyn Bevan, as well as to Dan Berliner, Sierd Hadley, Tom Hart, and Dan Honig. Thanks to Jessica Bergmann of the Education Commission for feedback and oversight of the production of this report, Debra Foulks for copyediting support, and Gloria Roh for graphic design.

All opinions and errors are the responsibility of the authors and should not be attributed to these institutions.

Please cite this paper as:

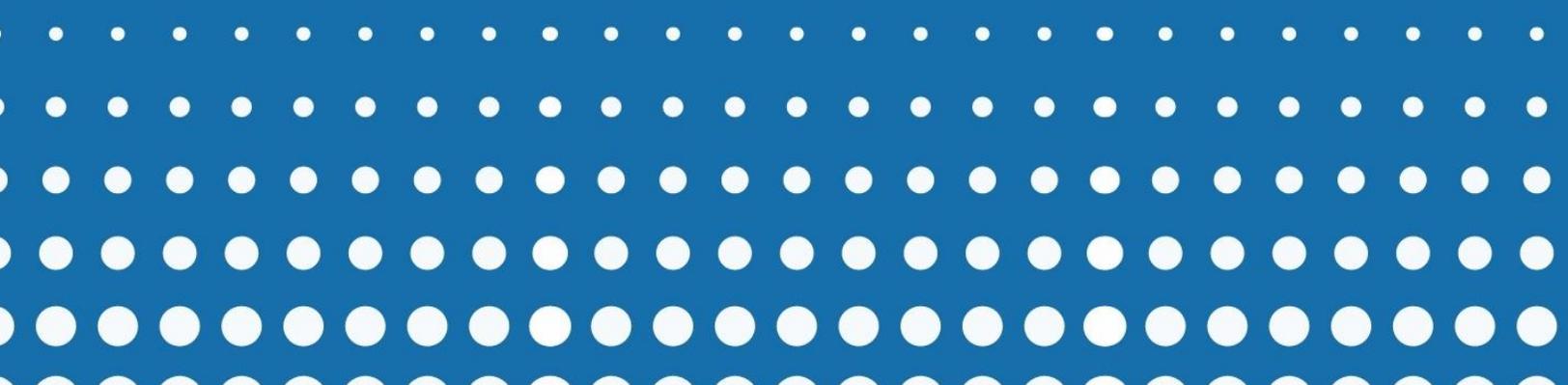
Mansoor, Zahra, Dana Qarout, Kate Anderson, Celeste Carano, Liah Yecalotecle, Veronika Dvorakova and Martin J. Williams. (2021, July). "A Global Mapping of Delivery Approaches." DeliverEd Initiative Working Paper. Education Commission and Blavatnik School of Government.

About DeliverEd

The DeliverEd Initiative was launched in 2019 and aims to strengthen the evidence base for how governments can achieve their policy priorities through delivery approaches. Globally, more than 80 countries have used delivery approaches to achieve better outcomes with reform and policy implementation. But to date, little empirical evidence, especially from low- and middle-income countries, exists on the effectiveness of delivery approaches and what design choices and contextual features contribute to this impact. The DeliverEd Initiative aims to fill this evidence gap and create a better understanding of what practices leaders can adopt to improve their policy delivery and reform efforts.

The Education Commission is leading this initiative in partnership with the Blavatnik School of Government, the University of Toronto, and with funding from the UK Foreign, Commonwealth and Development Office. Other partners include IDEAS, Georgetown University's [gui2de](#), and the World Bank.

For more information about DeliverEd and to view other related research and policy engagement materials, visit www.educationcommission.org/delivered-initiative.



A Global Mapping of Delivery Approaches

Zahra Mansoor, Dana Qarout, Kate Anderson, Celeste Carano, Liah Yecaló-Teclé, Veronika Dvorakova, and Martin J. Williams

July 2021

Abstract

The past 25 years has seen numerous governments worldwide adopting new institutional forms, such as delivery units and reform labs, to try to improve service delivery. We conduct a systematic global search and mapping of these *delivery approaches*. We identify 152 instances of delivery approach (DA) adoption from 80 different countries, ranging from the center of government down to provincial and local levels, with an accelerating trend of adoption since 2010. The majority of these include education as a focus sector. The main finding that emerges from our analysis is that there is no single model that characterizes the design of such approaches or the purposes for which they are adopted. However, we do identify a number of patterns, including that DAs in lower-income countries are more likely to focus mainly on output-type goals (as opposed to outcomes), use external consultants, and utilize accountability- and incentive-driven mechanisms relatively more heavily than their counterparts in middle- and high-income countries. While our findings are purely descriptive and cannot be used to infer whether DAs are effective, they give an indication to policymakers about the menu of potential options available to them in designing DAs. This global mapping can also provide useful context for existing and future country-level case studies that investigate the effectiveness of DAs and seek to explore which designs might be more effective in different contexts and for different purposes.

1 Introduction

A major trend in public administration reform over the last 25 years has been the creation of dedicated structures and processes to improve public service delivery across a priority sector or even the whole of government. These *delivery approaches* seek to support the implementation of priority policies or achieve priority outcomes by combining a set of managerial tools in a novel way to improve the functioning and performance of government bureaucracies. Such approaches encompass a range of institutional forms, from delivery units to reform labs and other institutional configurations. While the use of delivery approaches (DAs) is increasingly widespread, discussion around their potential benefits and limitations is often conducted with a limited understanding of the variation of patterns of their adoption and design worldwide. This risks unduly narrowing the view of such units for researchers and policymakers alike. This challenge is particularly acute in the education sector, where governments worldwide are increasingly seeking institutional mechanisms to improve education quality and deliver systemic reforms, and may benefit from a broader view of how other governments have approached this issue.

This paper addresses this adoption and design gap by creating a comprehensive database of delivery approaches worldwide, based on a systematic search, mapping, and coding protocol. In doing so, we build on a small but significant existing literature of reviews, how-to guides, and case studies focused on delivery units (e.g., Barber et al., 2010; Barber et al., 2011; Alessandro et al., 2014; Todd, 2014; Barber, 2015; Gold, 2014; Harrison, 2016; Gold, 2017; Delivery Associates, 2018; Lafuente & Gonzalez, 2018). Following Williams et al. (2020), we define a delivery approach (DA) as “an institutionalized unit or structured process within a government bureaucracy that aims to rapidly improve bureaucratic functioning and policy delivery by combining a set of managerial functions in a novel way to shift attention from inputs and processes to outputs and outcomes.” This definition includes not just delivery units but also related types of institutional and managerial approaches such as delivery/reform/innovation labs, policy monitoring or implementation units in executive or ministerial offices, and structured processes such as reform teams and committees to improving delivery. This gives our study a somewhat broader scope than much of this existing literature and allows us to better illustrate the range of institutional configurations available to governments. We perform our systematic search for 199 national governments worldwide (see list of countries in Appendix A) and also at subnational level for a nonrepresentative sample of six countries where such approaches are known to have been used.¹

¹These countries are Australia, Canada, China, India, Pakistan, and the United States.

Our global mapping codes cases of DAs adoption, along with a set of its key design features such as: the level of government at which the approach operates; the type of institutional structure; the nature of its goals; whether the approach involves donor support or technical assistance; the type of staffing structure; and the intended pathways (i.e., bundles of managerial functions) through which it seeks to achieve its impact. With respect to the latter, our coding follows the conceptual distinction in Williams et al. (2020) between two distinct pathways that DAs can draw on and combine to achieve their impact: one that relies on routines related to accountability and incentives (e.g., through establishing top-down targets, incentives, and controls) and another that establishes routines for collaborative problem-solving and organizational learning. Together, these variables (each presented with illustrative examples) allow us to paint a rich, descriptive picture of the many different ways in which DAs can be designed.

In total, we identify 152 cases of DAs being adopted across 80 countries. Of these, 107 cases are DAs being adopted by national governments within either particular sectors or across the whole of government, and a further 29 instances of adoption are at subnational levels (e.g., provinces, cities) within the selected number of countries for which we also undertook a systematic subnational search. Three cases identified were adopted at the frontline-level, and there was insufficient data to indicate the level at which the DA was set up for the remaining 13 cases. Of the 152 cases, 58% work across the whole of government or multiple sectors (approximately two-thirds of which cover the education sector) and 33% are focused on a specific sector (of which 40% are in the education sector), while there was insufficient data for 9% to indicate the sector of the DA. For simplicity, we focus our analysis on the 142 cases in our database where the DA takes the form of a single unit, directorate, lab, or other organizational structure and exclude 10 cases which match our definition of a DA but are not housed within a clearly defined organizational structure. We discuss this in greater detail below and later return to these cases to illustrate the wide range of options available to policymakers interested in improving delivery. Using a rich range of secondary data sources collected as part of the systematic mapping process, we also code a set of key characteristics about the goals, design, structure, staffing, resourcing, and function (i.e., types of managerial routines adopted) of each case to describe global patterns and trends in the design of DAs.

The key finding that emerges from our analysis of the design of DAs is that there is no single model or design that characterizes such approaches or the purposes for which they are adopted. Instead, there is substantial variation across every aspect of the goals, structure, and functions of these approaches. While this suggests that many of the factors driving the adoption and design

of DAs are fairly idiosyncratic and driven by contextual particularities, we do also observe some trends across countries. One such pattern relates to the use of delivery approaches in lower-income countries, which are more likely than DA in their middle- and high-income counterparts to focus mainly on output-type goals (as opposed to outcomes), use external consultants, and utilize accountability- and incentive-driven managerial mechanisms. While existing discussions of delivery units illustrate some of these variations in design and purpose (Alessandro et al., 2014; Todd, 2014; Gold, 2014; Harrison, 2016; Gold, 2017; Lafuente & Gonzalez, 2018), our data illustrate the global extent and scope of these variations. Governments considering adopting a DA, and the development partners or technical advisors supporting them, must therefore make a conscious set of decisions about the goals, structure, integration, and intended functions of their approach, and they have a very broad menu of options from which to choose.

In addition to describing variation along each of these dimensions in isolation, we also use simple statistical methods to understand whether certain design decisions tend to be correlated with each other. We find evidence that such patterns do exist. We find that DAs that use management routines that rely more heavily on accountability and incentives as levers for improvement tend to focus on delivering specific outputs or reforms, while approaches that rely more on routines involving problem-solving and organizational learning (or an equal mix of both accountability and problem-solving routines) tend to focus on achieving improvements in outcomes or a balanced combination of outcomes and outputs. Further, DAs that rely more on accountability and incentives routines are more likely to be housed within preexisting government units or structures, whereas DAs that employ routines involving problem-solving and organizational learning (or an equal mix of both accountability and problem-solving routines) tend to exist in units that have been newly established or reorganized. Finally, we also find that DAs that are housed within preexisting government structures also tend to recruit staff exclusively from within the civil service, whereas DAs that are housed in either new or reorganized structures tend to recruit entirely from outside the civil service or have mixed staff (i.e., a combination of civil servants and recruits from outside the civil service).

We relate these patterns to existing hypotheses and debates from the literature on delivery units and policy implementation as well as to practical insights on why these design choices tend to co-occur in practice (e.g., ACET, 2010; Barber, 2015; Khan et al., 2016; Gold, 2017). However, we emphasize that the fact that we observe these patterns worldwide does not imply that it is necessarily optimal for these design choices to be made together; we simply observe governments making them in practice. Further evidence from country-level studies that can

identify the performance implications of these design decisions, which is being undertaken as part of the broader DeliverEd Initiative, is necessary to make normative recommendations about what governments should do.

While our paper’s findings advance our understanding of the adoption and design of DAs worldwide, our analysis has several limitations. First, we aim merely to document empirical patterns in the adoption, design, and experience of delivery approaches worldwide, not to make claims about the effectiveness of these approaches or which design configurations are best. Rather, by systematically documenting cases, we aim to provide a fuller picture of countries’ efforts to improve bureaucratic functioning via institutional solutions like delivery units and reform labs and to identify common patterns that can help policymakers anticipate the choices and challenges they might face in adopting a DA in their context (as well as refine questions and hypotheses for further research).

Second, while our intensive and systematic search aimed to produce a comprehensive database of delivery approaches, it is almost inevitable that some cases were not possible to detect due to a lack of institutional or secondary documentation. This is particularly true of relatively small and more highly integrated initiatives, which are less likely to be documented via separate institutional websites, press releases, and so on. While our methodology describes the extensive efforts we made to minimize any bias resulting from differential data availability and to ensure consistency across countries, it is likely that some less-well-documented cases may exist that are not captured in the dataset. Our current global dataset is only a start to systematically mapping and capturing the key variations in DAs around the world; there may be scope in the future to expand research efforts and the database to provide a more comprehensive overview of global efforts aimed at improving implementation of public services for citizens.

Third, this paper focuses on gathering a wide breadth of data to establish general patterns and trends, but the necessary trade-off for achieving this breadth is relatively less depth on each case, especially as we restrict ourselves to using publicly available sources to ensure consistency of coding. While we do include qualitative information on some cases to illustrate our findings, and Appendix C contains an exhaustive coding protocol with specific examples, deeper and more nuanced discussion of specific cases will have to come through future primary data collection within DeliverEd and by other researchers.

Finally, although the scope of our dataset includes a wide range of structures and processes, governments adopt many strategies simultaneously to improve delivery, and not all of these fall

within our definition of a DA. There are, therefore, countless government efforts to improve service delivery that fall outside of our scope, so our paper is not a comprehensive review of efforts to improve service delivery, nor do we mean to imply that DAs are necessarily the most effective means to achieve such improvements.

The remainder of our paper proceeds as follows. Section 2 discusses key definitions and the scope of our review as well as highlights selected existing reviews and related literature, and Section 3 lays out our methodology for the systematic search, mapping, and coding of cases. Section 4 then presents our descriptive findings about patterns in the adoption of DAs worldwide, and Section 5 presents descriptive findings about the design features of these approaches. Section 6 discusses the observed statistical relationships among these different design features, and Section 7 briefly discusses a selected set of strategies and structured processes governments used to try to improve delivery that do not meet all the criteria for inclusion in our main analysis but nonetheless help illustrate the wide range of options available to governments considering adopting a DA. Section 8 concludes by reflecting on the findings and limitations of this review, as well as their implications for policymakers.

2 Definition and Scope

The scope of our global review is delimited by our definition of a *delivery approach* (DA) as “an institutionalized unit or structured process within a government bureaucracy that aims to rapidly improve bureaucratic functioning and policy delivery by combining a set of managerial functions in a novel way to shift attention from inputs and processes to outputs and outcomes” (Williams et al., 2020). Such approaches have rapidly proliferated worldwide over the past 2 decades as policymakers have often sought to improve the implementation of policy priorities by government bureaucracies, with the New York City ‘CompStat’ policing program and the United Kingdom’s Prime Minister’s Delivery Unit as initial archetypes in the 1990s and early 2000s (Gold, 2017). This proliferation includes the education sector, where the establishment of a DA has become increasingly used as a tool for governments and development partners seeking system improvements (e.g., Andrews, 2014; Harrison, 2016; Lafuente & Gonzalez, 2018; FCDO, 2019).

The definition above makes clear some of the features that such approaches share and which

distinguish them from adjacent concepts, strategies, and factors that also affect policy implementation and service delivery outcomes - as nearly all aspects of a government bureaucracy ultimately do. In particular, DAs are novel in their context, relative to the status quo bureaucratic set-up; operate primarily through the exercise of managerial tools internal to a bureaucracy to change bureaucratic behavior; and aim to focus attention and effort on achieving targeted outcomes and outputs at the end of the delivery chain, rather than just on controlling resource inputs or following regulations. While most policy and academic attention has focused on *delivery units* (which are themselves a diverse category) as exemplars of DAs (e.g., Barber et al., 2011; Barber, 2015; Gold, 2014; Harrison, 2016; Kohli et al., 2016; Gold, 2017; Lafuente & Gonzalez, 2018), this inclusive definition also encompasses a broader set of institutional forms such as delivery/reform/innovation labs, policy monitoring or implementation units in executive or ministerial offices, and reform teams (if intended to be institutionalized and continuous, rather than to execute a specific project). The definition also includes selected structured processes that do not have a specific organizational form (i.e., are not a defined “unit”); however, for practical reasons explained in Section 3, we exclude these from our main empirical analysis and discuss them in Section 7. Finally, DAs can in principle be located at any level of government, from the center of government (CoG) to national-level ministries, subnational governments, or even frontline organizations such as schools, which we discuss in detail in Section 5.

While this definition is deliberately inclusive of a range of institutional forms that share a common goal and key features, we follow Williams et al. (2020) in excluding from our global review a range of other policy tools that governments may use to try to improve outcomes in the education sector and beyond. As DAs (by our definition) are spearheaded by government or carried out in collaboration with government, we exclude initiatives that are externally led by non-governmental organizations (NGOs), development partners, or local communities and service beneficiaries (though these groups may be involved as stakeholders or co-producers in a DA). For instance, community-based approaches (e.g., community monitoring of schools) may be used by governments to try to improve education outcomes, but are excluded from our definition - although community members may play important supporting roles in a DA. We also exclude initiatives based solely around financing models such as results-based financing, pay-for-performance, public-private partnerships, and social impact bonds, as well as efforts to improve performance through the allocation of additional resources. In addition, we consider DAs to be institutionalized and intended to be continuous in nature. This excludes initiatives such as one-off training interventions, cash transfers, units set up to deliver the goals of a specific project (e.g., implementation units for donor-funded programs) or other discrete management reforms.

However, we include DAs that are institutionalized with the objective of existing for a relatively short time span; we return to the question of the time horizon of DAs in our empirical review below.

This paper builds on an important existing body of policy-oriented literature, as well as relevant insights from past government and public service delivery experience among DeliverEd’s High-Level Advisory Group (HLAG) members. Several excellent multi-country reviews have drawn on practitioner experience, case studies, and policy surveys to present a rich picture of the use of delivery units worldwide, in the education sector and more broadly (e.g. Andrews, 2014; Todd, 2014; Shostak et al., 2014; Gold, 2014; Barber, 2015; Kohli & Moody, 2016; Harrison, 2016; Gold, 2017; Delivery Associates, 2018; Todd, 2018; Todd & Waistell, forthcoming;). In particular, past efforts by Gold (2017) and Todd and Waistell (forthcoming) to map delivery units around the world have provided a wealth of qualitative information about the management functions these units leveraged to focus bureaucrats on improving implementation and about the challenges these units faced in achieving tangible results for citizens. We build on these studies by adopting a slightly broader scope of the institutional tools available to policymakers (focusing on DAs rather than just delivery units), updating and systematizing the mapping of cases worldwide, coding the cases based on the Williams et al. (2020) conceptual framework, and extending their analysis of typologies and patterns. There also exist numerous case studies of specific instances of delivery unit use (e.g., Scharff, 2012; Iyer, 2012; Simson, 2013; Freeguard & Gold, 2015; CPI, 2016; World Bank, 2017; Delivery Associates, 2018; Lafuente & Gonzalez, 2018). While we draw on these in our review to support case identification and coding and for selected qualitative examples, we aim to complement these in-depth case studies through the breadth of this global review.

Our paper also links to a broad range of academic literature on bureaucratic performance and service delivery, too numerous to enumerate fully here. These linkages include: the diverse literature on public service delivery which offers a range of theoretical perspectives on patterns in implementation failures (e.g., World Bank, 2004; Brinkerhoff, 2004; Hood & Dixon, 2010; Chamber et al., 2012; Pritchett, 2015); the literature on the use of performance information and performance management techniques, mainly focused on governments in high-income countries (e.g., Bevan & Hood, 2006; Radin, 2006; Moynihan, 2008; Moynihan et al., 2008; Hammerschmid et al., 2016; Gerrish, 2016); and the literature on patterns of public sector reform since the 1990s (e.g, Pollitt & Bouckaert, 2004; Andrews, 2013; Goldfinch et al., 2013; Hammerschmid et al., 2016). While this paper’s main contribution is as an empirical review of practice, and we do not

seek to generate new theoretical insights or make practice-oriented recommendations, we link our empirical findings to these and other literature.

3 Search and Coding Methodology

3.1 Search

To identify cases of DA use, we conducted a systematic desk-based search covering 199 countries.² Our search focused mainly on DAs located within national governments because time constraints made it infeasible to conduct systematic searches of subnational jurisdictions for all countries. However, we did conduct systematic searches of subnational jurisdictions for six countries: Australia, Canada, China, India, Pakistan, and the United States. Where we were able to identify subnational cases in other countries, we also included these in our database. While our list of DAs used at subnational level is, therefore, not as comprehensive as our national-level list, nor necessarily representative of the use of DAs at subnational levels nationwide, it nonetheless allows us to provide some insight into patterns of use and design at the level of cities and provinces. We defined the start period of our search as 1990, several years before the establishment of the first self-styled delivery units, and the end in 2020.

For each country or subnational jurisdiction, our search protocol drew on a wide range of potential sources: existing lists of delivery units from prior reviews (e.g., Alessandro et al., 2014; Harrison, 2016; Gold, 2017; Lafuente & Gonzales, 2018) and cases that had been raised in meetings with DeliverEd stakeholders; databases of reforms, donor projects, and country case studies from sources such as the AidData donor projects database, the World Bank’s project repository, the Global Delivery Initiative database, the Inter-American Development Bank and the African Development Bank databases, and the Institute for State Effectiveness Reform Sequencing Tracker; systematic keyword searches of internet search engines (Google and Google Scholar), as well as the University of Oxford’s SOLO journal index; and targeted searches based on references to potential DAs from other sources. This exhaustive process enabled us to build on existing lists of DAs and identify previously unrecognized potential cases, including those that refer to themselves in non-“delivery” terms, such as “monitoring and evaluation units” and

²The full set of countries is listed in Appendix A and excludes occupied or disputed territories and constituent territories of the United States, the United Kingdom, France, and the Netherlands.

“reform secretariats”. The full search protocol is contained in Appendix B.

For each identified potential case, we first collected as much publicly available information on the case as possible (from academic, policy, and other sources), then evaluated whether the case matched our definition of a DA given in Section 2 above. Monitoring and Evaluation Units or Implementation Units that were set up to track the implementation of donor-funded projects that did not fit in with a broader national strategy were excluded from the sample. For example, Project Implementation Units (PIUs) funded and staffed by donor agencies for the narrow purpose of delivering a specific project were not considered DAs.³ In total, we identified 152 cases (at both national and subnational levels) from 80 countries as meeting our definition of a DA.

Our definition of DAs covers both discrete organizational units, as well as the more amorphous category of structured intra-governmental processes that aim to improve delivery but do not have a specific unit leading them. However, the nature of these processes meant that there was much less documentation publicly available about them, and it was, therefore, more difficult to comprehensively identify and code them. Hence, we excluded the 10 cases of structured-process DAs that we identified from our analysis in this paper, focusing instead on the 142 (93%) cases that took the form of a specific unit and for which our global database is more comprehensive. We return in Section 7 to discuss the structured-process DAs and illustrate to readers how these are used and designed.

3.2 Coding

After identifying these cases, we then coded a set of variables about the key characteristics and design of each case based on the publicly available data collected on each case. These variables can be divided into two sets: (1) basic characteristics of the case’s name, context, sectors covered, and start and end dates; and (2) variables related to the goals, design, and functions of the DA. This latter set of variables are coded based on the conceptual framework laid out in Williams et al. (2020). We discuss the coding of these variables and their interpretation in greater detail as we present descriptive statistics on them in Sections 4-5. Appendix C contains the full codebook used, as well as examples of how cases were coded for various variables, illustrating how the

³PIUs were identified in the OECD database on Overseas Development Assistance (ODA) covering data collected for the Survey on Monitoring the Paris Declaration.

information available was interpreted to inform coding decisions.

Coding was conducted by this paper’s authors. We first piloted the coding scheme on a sample of 17 countries before coding the full set of cases. To ensure consistency in coding, we held weekly discussions to converge on the interpretation of information to align with and clarify our conceptual definitions of the various constructs. Where there was ambiguity in the information about specific cases, two coders worked on coding the case in question. For a small number of variables for which interpretations had diverged between coders, we recoded to ensure consistency in the final dataset. In the full coding process, each case was coded by a single coder based on the available evidence about each case, with frequent discussions among coders and scrutiny over ambiguous cases throughout the coding process to harmonize interpretation. Our aim was thus to achieve consistency of coding across cases. For each variable, we give illustrative qualitative examples to indicate the interpretations we applied throughout the coding process. The statistics and variable definitions presented in this paper are thus consistent across cases.

Our coding relied on publicly available information, as the large number of cases made it infeasible to conduct primary data collection in a consistent manner across all cases. This meant that the volume and depth of information available differed across cases and that different sources sometimes gave different information or differed in their interpretation. We, therefore, triangulated across these various sources of information to code each case as accurately and consistently as possible. Where there was no public information, insufficient level of detail, or conflicting information on a variable, it was coded as “unclear/don’t know”⁴. The reliance on public information also means that while we tried to code each variable on a *de facto* basis where possible (i.e., how it actually functioned), many cases had to be coded based mainly on official sources that tend to focus more on *de jure* (i.e., how it was supposed to function) features and practices. Finally, for some cases that existed for longer periods of time, the goals and/or design of the case changed over time. Where a variable could have been coded differently in different time periods, we coded it according to what pertained during the initial phase or years of the case. Given these data limitations, we focus our analysis on overall patterns and trends rather than on specific cases.

⁴Given that data availability varied across variables, our descriptive trends for each variable rely on DAs for which there was available data for the variable in question and the sample size thus differs across variables. While we have no indication that data availability is correlated with any of the characteristics we are coding, we nevertheless emphasize that the descriptive patterns we present cover only those cases with sufficient data availability to code the variable in question.

4 Adoption of Delivery Approaches

This section provides a descriptive summary of the patterns in the adoption of DAs that we observe in the data, presenting the distribution of the cases across regions, country income levels, and trends in adoption over the last couple of decades. We observe two main trends in adoption:

- The adoption of DAs is increasingly common. DAs were adopted in 40% of countries in the past 3 decades at either national or subnational level, and more than half of DAs were initiated after 2010.
- DAs have been adopted in all regions of the world, with the largest number of cases identified in Sub-Saharan Africa and in Latin America and the Caribbean. Roughly one third of cases are in high-income countries (31%), followed by upper middle-income countries (28%), lower middle-income countries (23%), and finally low-income countries (18%).

The remainder of this section presents these trends in more detail and gives further information about their coding, using selected examples for illustration.⁵

Based on our definition of a DA, we were able to identify 142 cases of adoption around the world.⁶ Out of the 199 countries we reviewed, DAs had been adopted at either a national or subnational level in 80 countries (40%).⁷ This estimate is significantly higher than estimates from other recent reviews by Gold (2014, 2017) and Todd and Waistell (forthcoming), which uncovered 44 cases across 33 countries. This is in part due to our broader focus on DAs rather than just delivery units and also due to the increased adoption of DAs over the last 3 years. Figure 1 below presents the adoption of DAs over time and highlights that the rapid proliferation of these approaches is a relatively recent trend, with the popularity particularly surging in the last decade.⁸ The novelty

⁵The descriptive analysis and figures that we present exclude missing observations for which adequate data to code each variable was not available.

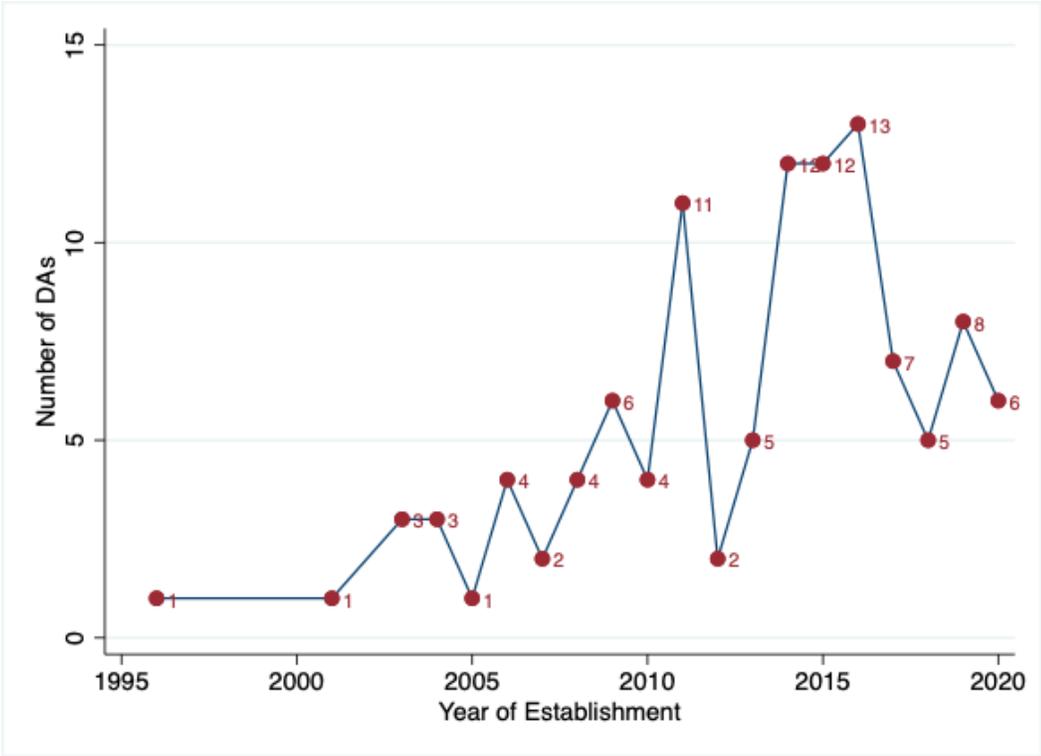
⁶An additional 10 cases were identified that were not institutionalized units but rather structured processes that carried out institutional and managerial approaches that are similar to such units; these cases are excluded from all figures and statistics in Sections 4-6.

⁷Excluding structured process-type DAs (as we do in the remainder of Sections 4-6), DAs have been adopted in 73 countries.

⁸While this historical trend is indicative of the rising popularity of DAs in the last 10 years, our mapping is based on the availability of public data through secondary sources, so it is possible that our review is more comprehensive of relatively recent cases of DAs adopted around the world compared to earlier cases.

of this global phenomenon has also been noted in other case studies on delivery units around the world, including Gold’s (2017) *Tracking Delivery: Global Trends and Warning Signals* report in which the author notes that over half of the 31 national and regional delivery units identified across 25 countries were established between 2015 and 2017.

Figure 1: DELIVERY APPROACHES HAVE BECOME A GLOBAL TREND IN THE LAST DECADE



Note. $N = 113$.

Figures 2 and 3 show the number of DAs by region. Just under half of DAs were identified in Sub-Saharan Africa and Latin America and the Caribbean. The highest number of cases was found in Sub-Saharan Africa, where 39 distinct cases were identified across 21 countries, making up just over a quarter of total cases (27%) in our database. Following Sub-Saharan Africa, Latin America and the Caribbean have the second largest number of DAs, reaching 29 cases in 16 countries, which makes up a fifth (20%) of the total number of cases identified in the mapping. We found only 9 cases each in North America and the Middle East and North Africa (MENA), the smallest number of DAs in any region. Approximately one third of the cases are in high-

Figure 2: SUB-SAHARAN AFRICA HAS THE LARGEST NUMBER OF DELIVERY APPROACHES WORLDWIDE, FOLLOWED BY LATIN AMERICA & THE CARIBBEAN

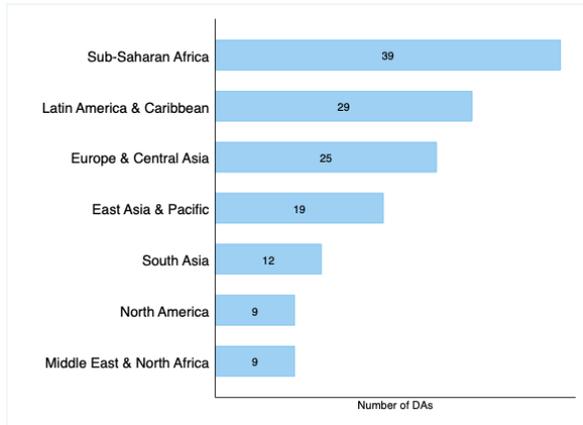
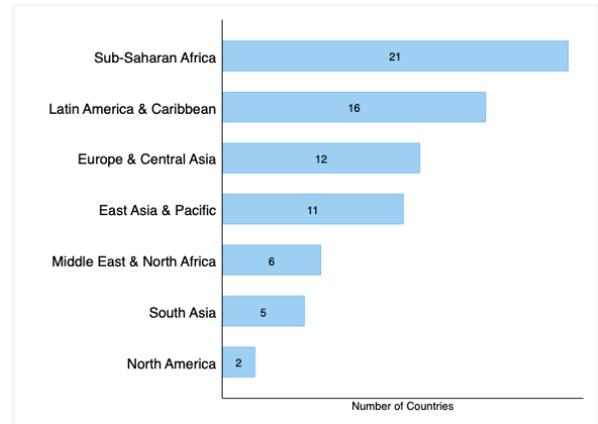


Figure 3: THE LARGEST NUMBER OF COUNTRIES WITH AT LEAST ONE DELIVERY APPROACH CAN BE FOUND IN SUB-SAHARAN AFRICA AND LATIN AMERICA AND THE CARIBBEAN



income countries (31%) followed by upper middle-income countries (28%), lower- middle income countries (23%), and finally low-income countries (18%).

5 Designing Delivery Approaches

5.1 How are delivery approaches structured and organized?

In this subsection, we discuss the trends observed in the design/set-up of DAs, including the level of administration at which they operate, the sectors they cover, the staffing structure they have adopted, and the institutional arrangement that hosts the DA's mandate and function. Furthermore, this subsection includes a description of the technical and funding assistance provided to DAs by external entities, including international NGOs, donors, and private consulting firms. We observe four main trends in the design of DAs along the aforementioned features worldwide:

- The majority of DAs operating at national-level are located in the CoG (e.g., a prime min-

ister’s office) rather than within sector ministries, but the adoption of DAs at subnational levels is also increasingly common.

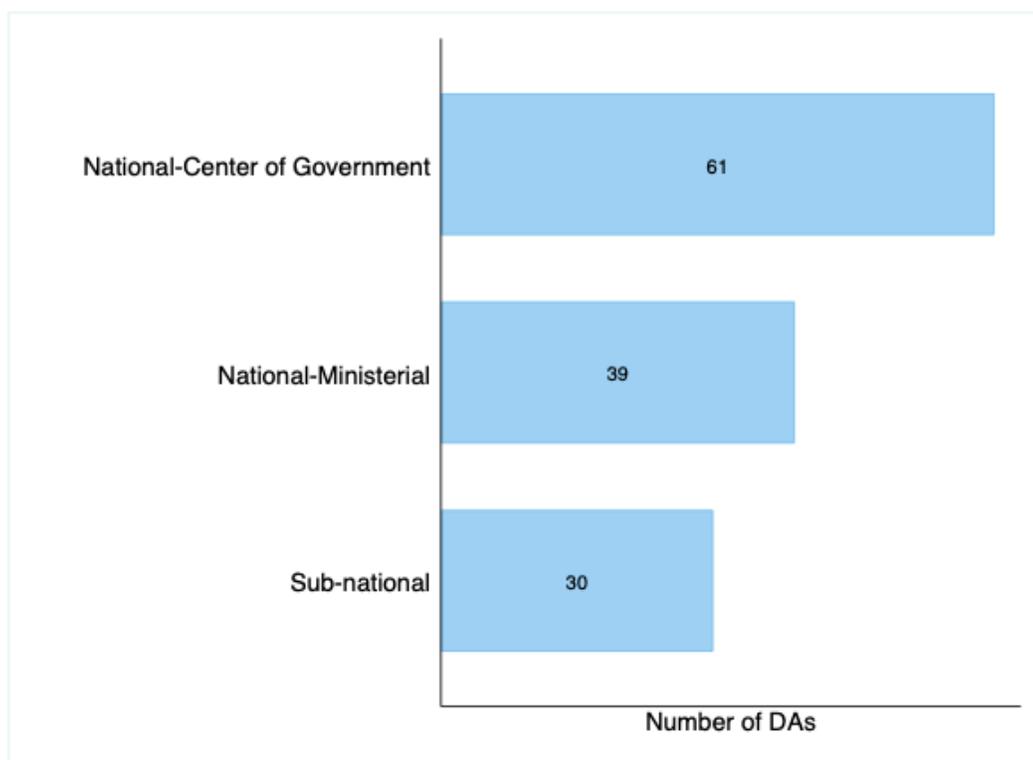
- The majority of DAs operate across multiple sectors rather than focusing on a single sector, but one third operate within a single sector. Forty-seven (47) percent of DAs include an education focus, either as a single focus sector or as part of a multisector DA.
- Most DAs (52%) were staffed exclusively by existing civil servants, although use of external recruitment for core staffing was also common, especially in middle- and low-income countries as compared to high-income countries.
- Under half of DAs (46%) received external funding (e.g., support from development partners or international organizations) for their set-up and/or functioning. Use of external technical assistance was more common (69% of cases), although far from universal in most regions.

With respect to the location within government where the DAs are established, we observe that national-level DAs are more likely to be set up within the President’s or Prime Minister’s Office (or equivalent) than in line ministries. Just under half of the DAs in our dataset (47%) were established within the CoG at the national level. Examples of such approaches include the PMDUs in the UK and in Tunisia, which are hosted in the Prime Ministry. Similarly, Sierra Leone’s Strategy and Policy Unit (SPU) is hosted within the President’s Office, delivering on the Presidential Agenda (Scharff, 2012). Since existing research on delivery units by Gold (2014, 2017), Harrison (2016), Lafuente & Gonzalez (2018), and Barber (2010, 2011, 2013, 2015) have highlighted coordination and convergence in governmental efforts as one of the main attractions of a DA to policymakers, it is not surprising that most cases in our dataset are at the central level where the approach can effectively manage cross-sectoral and multistakeholder coordination. Furthermore, since DAs have typically been leveraged to signal political urgency of certain priority areas (Gold, 2017; Lafuente & Gonzalez, 2018), a central level DA is perhaps more likely to allow for close proximity to the political sponsor. That said, we also find a considerable number of cases (39) set up within national-level ministries (30% of all cases). Examples of ministerial DAs include Ethiopia’s Education Sector Delivery Unit, Kosovo’s Central Delivery Unit within the Ministry of Finance, and Macedonia’s Transport Infrastructure Delivery Unit. We also documented another 30 cases (23%) at subnational levels such as district, provincial, state, city, and even sub-city administrations, both within the subnational equivalents of CoG offices

(e.g. mayor’s offices) and in subnational sector ministries.⁹ For example, the city of Baltimore in the state of Maryland in the United States established the Mayor’s Office of Performance and Innovation in 2019 and Brazil’s Sao Paulo Municipality set up its own delivery unit to monitor local services and priorities (Freeguard & Gold, 2015; Allesandro, Lafuente & Santiso, 2014).

⁹In our coding scheme (Appendix B) we included a category for ‘Frontline’ approaches, in which DAs operate at the level at which services are delivered, such as in schools or local councils. However, for feasibility reasons our search was not comprehensive at this level. We identified only two examples of such units, in the London Boroughs of Barking and Dagenham and Haringey (LBBD 2017; Etheridge & Thomas, 2015). For simplicity, we group these together with subnational units in our figures.

Figure 4: MOST CASES ARE ESTABLISHED AT THE NATIONAL LEVEL OF GOVERNMENT



Notes. $N = 130$, the figure above shows the distribution of identified delivery approaches across levels of government where they were set up. National- Center-of-Government (CoG) refers to central administration at the national level, typically the President's or Prime Minister's Office (or equivalent). National- Ministerial DAs are set up within line ministries to focus on single sector priorities at the national level. Subnational DAs refer to approaches set up at the district, provincial, state, city, or sub-city level. These cases could operate within the central administrative offices of these subnational units, but they differ from CoG DAs in that they do not function at the uppermost administrative level in government that run from the office of the Head of State. Subnational approaches were not searched systematically for all countries.

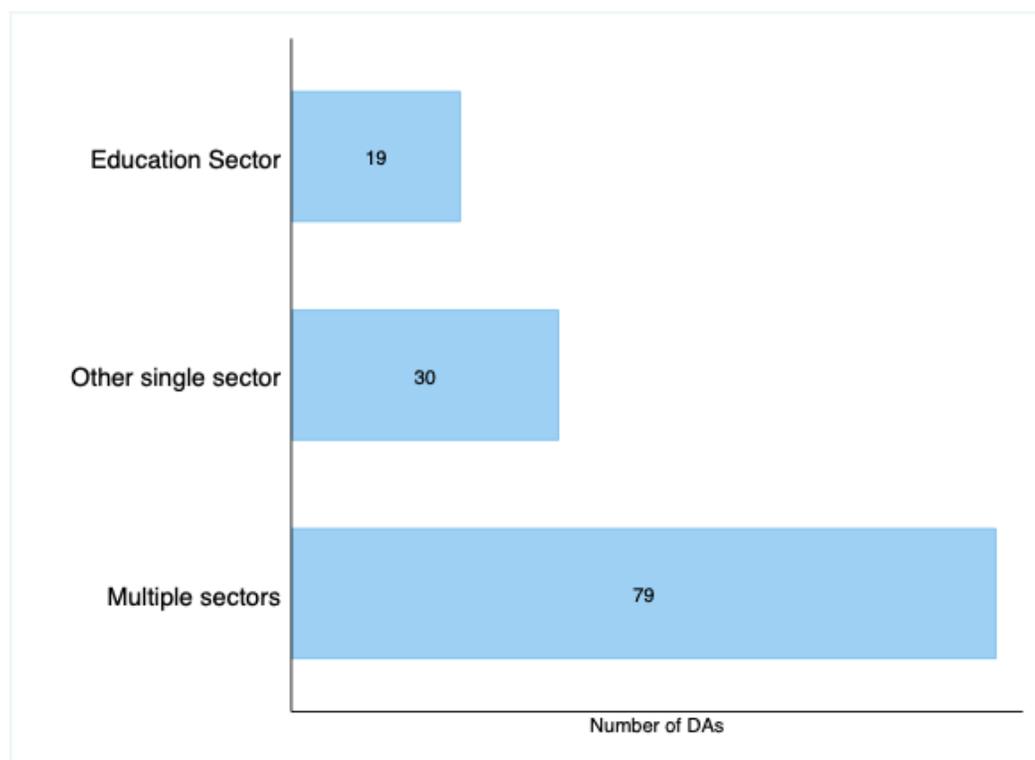
While we code DAs according to the level of government at which they are located or institutionalized, it is important to note that DAs established at one level often also interact with other levels of government in order to establish and leverage management routines across the delivery chain. For example, the Central Level Project Implementation Unit in Nepal is set up at the central level to monitor reform efforts within the education sector; however, it also operates across districts, creating coordination and planning mechanisms to ensure that the delivery

plan encompasses agents working directly with schools (MoE, Nepal, n.d.). Similarly, the DA in Oman is set up within the royal court but relies on a network of satellite structures across line ministries responsible for the implementation of national reform priorities (ISFU, n.d.). While we do not attempt to code such relationships in our cross-national mapping, they nevertheless constitute an important part of the functioning of DAs and are an important topic for future country-level work within the DeliverEd project.

Most DAs (62%) have a remit that includes multiple sectors across government, rather than focusing more narrowly on single sectors. Figure 5 illustrates this. In some cases these DAs aim to cover the whole of government, while in other cases they focus on a handful of priority sectors. This likely reflects the concentration of DAs in executive offices in the CoG and at subnational levels due to their role in amplifying executive oversight, as well as the aim of many DAs to improve cross-sectoral collaboration (e.g., Gold 2014; Todd & Waistell 2014; Lafuente & Gonzalez 2018; Allesandro, et al., 2014; Barber 2015; Harrison, 2016; Gold, 2017). The remaining 38% of DAs focus on single sectors. Cameroon's and Rwanda's Ministries of Agriculture, for example, have their own delivery units; the UK, Ireland, and Rwanda have adopted delivery units specific to the health sector; and Macedonia, Singapore, Liberia, and Kosovo have all adopted DAs to specifically monitor infrastructural reform. The education sector is perhaps the sector most frequently covered by DAs: 39% of all single-sector DAs are in the education sector, and 61% of multisector DAs for which adequate data exists include education among their focus sectors.¹⁰

¹⁰Our search process may have been slightly more likely to identify sector-level DAs within the education sector, since we were alerted to some cases through stakeholder outreach and most of our outreach was with education-sector stakeholders.

Figure 5: MOST DAs' REMITS ARE CROSS-SECTORAL



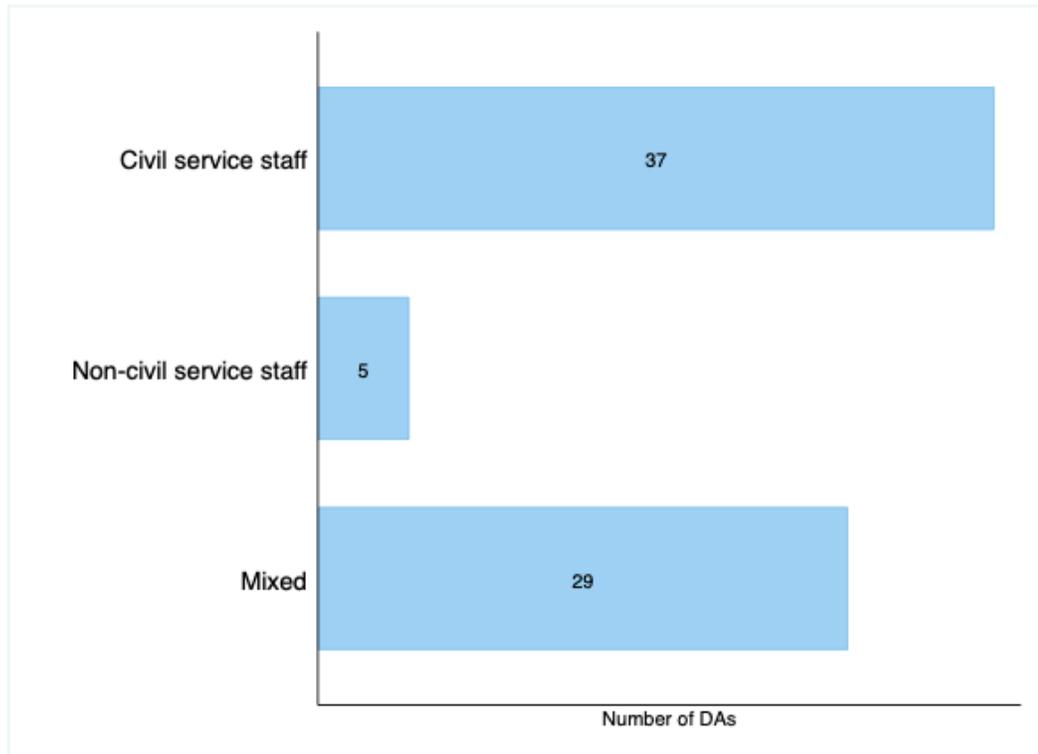
Note. $N = 128$

One frequent area of discussion with respect to delivery units is the question of how, and by whom, they are staffed (e.g. Barber 2015; Gold, 2017; Delivery Associates, 2018). Figure 6 shows that most DAs (52%) are fully staffed by core civil servants. However, the number of cases with mixed staff of both civil servants and external consultants, advisors, or new hires is quite substantial (41%), and only a handful (7%) of cases are staffed entirely with non-civil servants.

However, Figure 7 shows that these patterns differ substantially across income groupings, with the vast majority of DAs in low-income countries relying on human resources outside the civil service to function. Over two-thirds of DAs (69%) in low-income countries are staffed both from within the civil service and externally. Similarly, over half of cases in lower middle-income countries rely on talent found outside the civil service, while 88% of cases in high-income countries rely entirely on core public sector employees to implement the DA. The reliance of DAs in low-income countries on human resources outside the civil service could be due to the relative lack

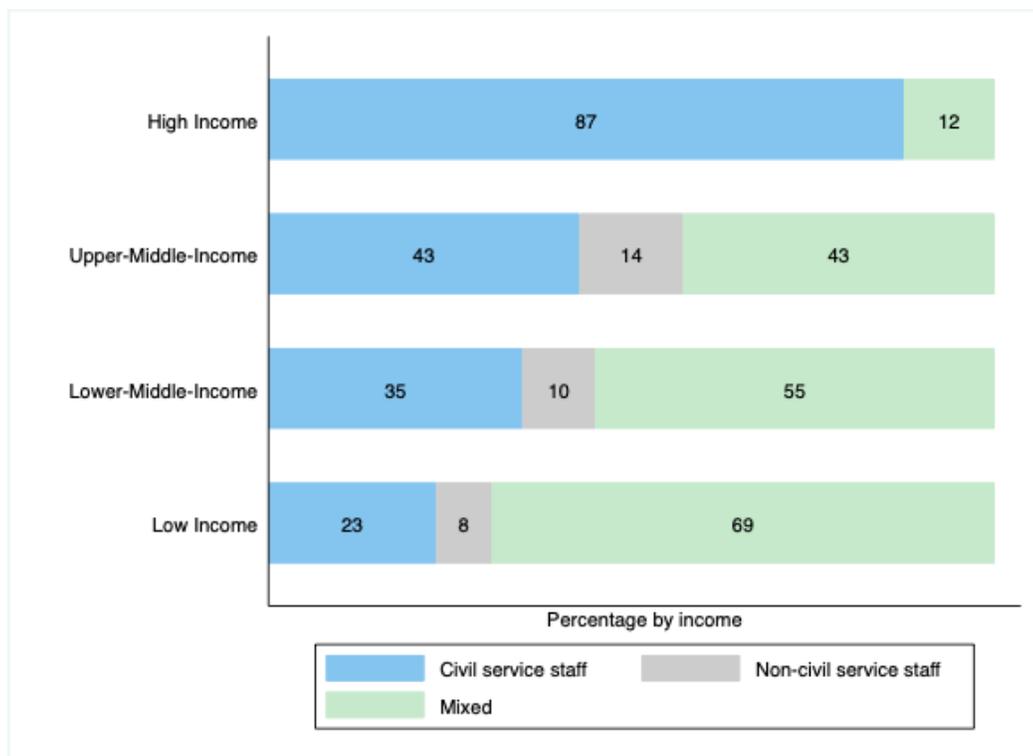
of requisite skills of existing civil service staff, as well as to the frequent involvement of external donors and technical advisors.

Figure 6: MOST DELIVERY APPROACHES ARE STAFFED EXCLUSIVELY WITH CIVIL SERVANTS



N = 71. Staffing was coded based on whether team members carrying out the delivery approach were entirely hired from within the civil service or externally. If the staff constituted both civil servants and consultants hired from the private sector or seconded from other organizations, the staffing feature of that particular case would be considered 'mixed'.

Figure 7: DELIVERY APPROACHES IN LOW-INCOME COUNTRIES ARE MORE LIKELY TO MAKE USE OF EXTERNAL CONSULTANTS OR ADVISORS

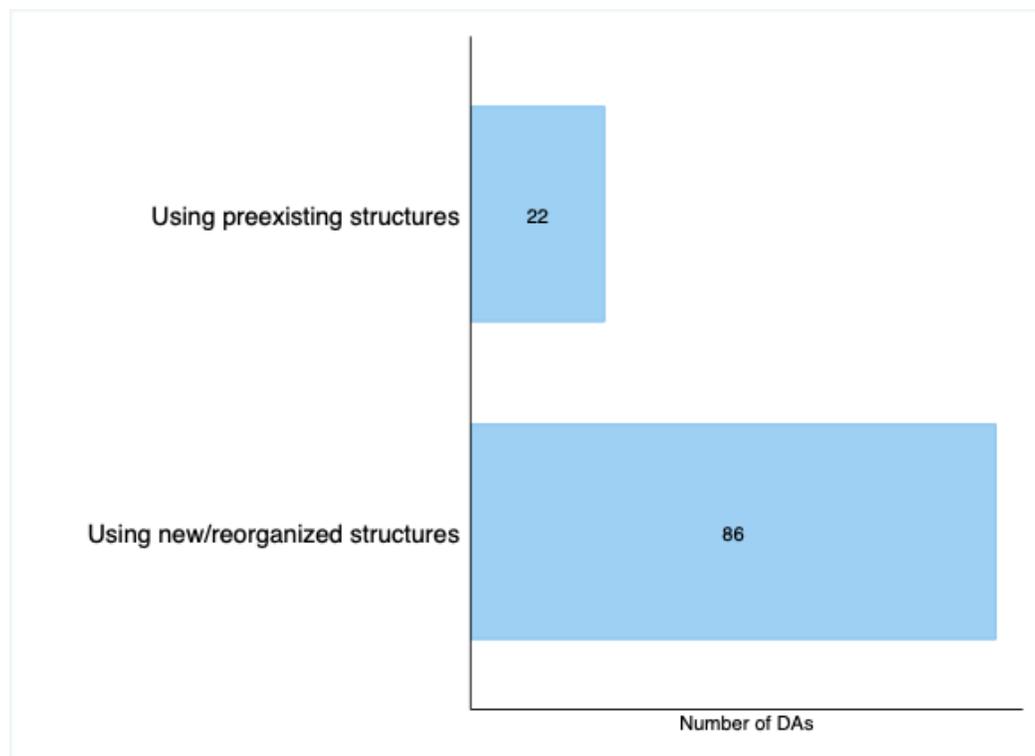


N = 71. Staffing was coded based on whether team members carrying out the delivery approach were entirely hired from within the civil service or externally. If the staff constituted both civil servants and consultants hired from the private sector or seconded from other organizations, the staffing feature of that particular case would be considered 'mixed'.

Another important distinction is the extent to which DAs are housed within preexisting bureaucratic structures and processes, as opposed to being inserted in bureaucracies as new or reorganized units. In the former case, adopting a DA means assigning new functions or responsibilities to an existing unit or team without any change to the overall 'org chart', while in the latter case, adopting a DA implies a significant restructuring of responsibilities and often the creation of a new unit that sits astride the preexisting bureaucracy. See Appendix B for full definitions and details of coding. We find that the vast majority (80%) of DAs fall into the latter category, while only 20% are housed within preexisting structures. The Implementation Support and Follow-Up Unit (ISFU) in Oman is an example of an approach that was hosted in a new bureaucratic structure because it was established as an entirely new unit that was created

through legislation and given the same status as a line ministry (ISFU, 2018). DAs that are housed within preexisting bureaucratic structures, on the other hand, operate within the context of long-established departments such as monitoring, coordination, or implementation units. For example, the Coordinación General De Secretaría within the Ministry of Education in Ecuador is fully embedded with the organizational make-up, reporting directly to the Minister of Education and overseeing the prioritization, resourcing, and implementation challenges of reform in the sector (MoE, *Ecuador*, n.d.). We find that DAs in high- and upper middle-income countries are more likely to have DAs that are fully housed within preexisting bureaucratic structures and processes, while none of the cases in low-income countries or in Sub-Saharan Africa are fully housed within the existing bureaucracy but are instead created as new/hybrid/reorganized units.

Figure 8: MOST DELIVERY APPROACHES ARE HOUSED IN NEW OR RE-ORGANIZED BUREAUCRATIC STRUCTURES



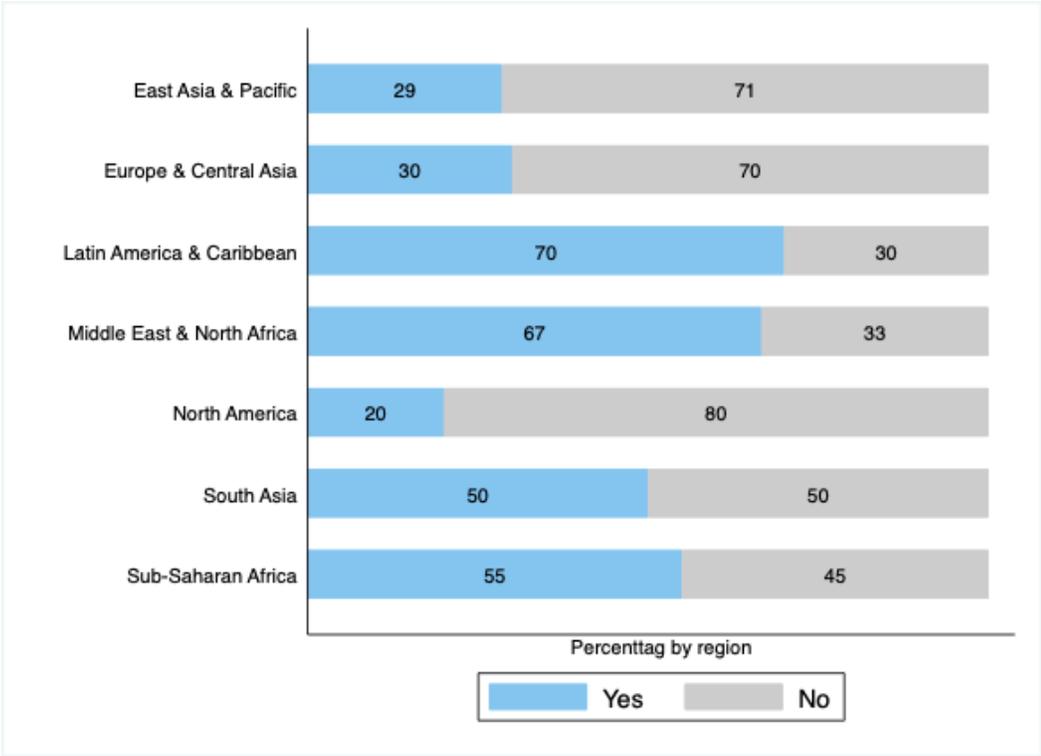
N = 108. This figure shows the variation across DAs in their institutional set-up. We distinguish between two types of structural arrangements: (1) DAs that are housed in or rely on preexisting structures/processes to implement the approach, and (2) DAs that leverage some innovation in the organizational and procedural make-up of the agency to develop and operationalize the new DA (for instance, setting up a new department or merging several departments and reorienting their functions and tasks to accommodate the new DA).

Many DAs worldwide receive external funding and/or external technical assistance. Figures 9 and 10 show the proportion of DAs in different regions and income groupings that we were able to identify as receiving external funding, while Figures 11 and 12 show the proportion we identified as receiving external technical assistance.¹¹ Receiving external technical assistance is

¹¹It is possible that some cases received external funding and/or technical assistance that was not disclosed in the public sources on which we relied. Where we were able to find information on these variables and no mention of external support was made, we coded these variables as 'No', and where there was inadequate information available to make a reasonable judgment about whether external report was received, we coded these variables as missing (and thus excluded them from the figures presented here).

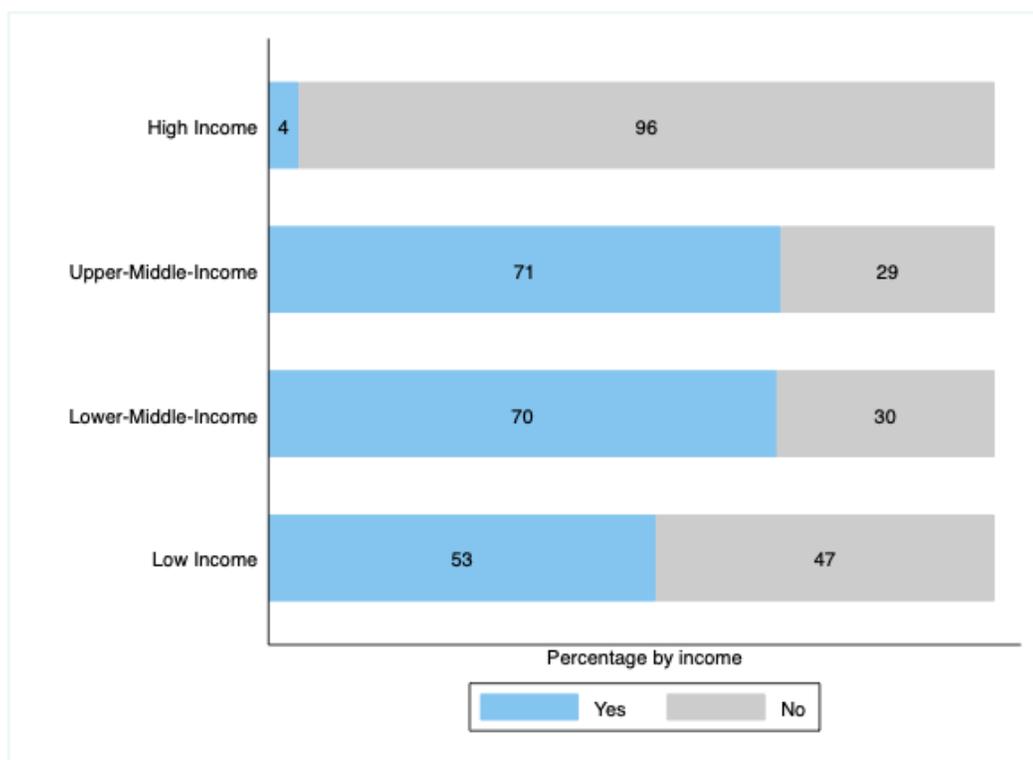
more common than receiving external funding, with 69% of all DAs worldwide receiving technical assistance and 54% receiving external funding. Unsurprisingly, both types of external support are generally more common in lower-income countries and regions, but even in high-income countries, 40% of DAs received some form of external technical assistance from consultants or other technical advisors. Examples of DAs that received external funding, such as from international donor banks, include Tanzania's President's Delivery Bureau, receiving funding from the World Bank, the UK Foreign, Commonwealth, and Development Office (FCDO), and USAID, and Kenya's Presidential Delivery Unit, with funding from the African Development Bank Group (ADBG). The majority of DAs in the MENA region across 6 countries have received some form of technical assistance, but less than half, mostly in upper middle- and lower middle-income countries, received external funding, including Tunisia, Jordan, Libya, and Morocco.

Figure 9: DAs IN REGIONS WITH MANY LOW- TO MIDDLE-INCOME COUNTRIES ARE MORE LIKELY TO RECEIVE EXTERNAL FUNDING



N = 76. External funding for a DA refers to receiving funding from donors, INGOs, private firms or other organizations outside national government for the establishment, design, and/or operationalization of the approach. This encompasses funding to set up structures and processes, salaries to compensate staff, funding or in-kind services for technical assistance for the design of the approach and consulting on best practices, and other financial support required to carry out the approach.

Figure 10: DAs IN HIGH INCOME COUNTRIES ARE LESS LIKELY TO RECEIVE EXTERNAL FUNDING

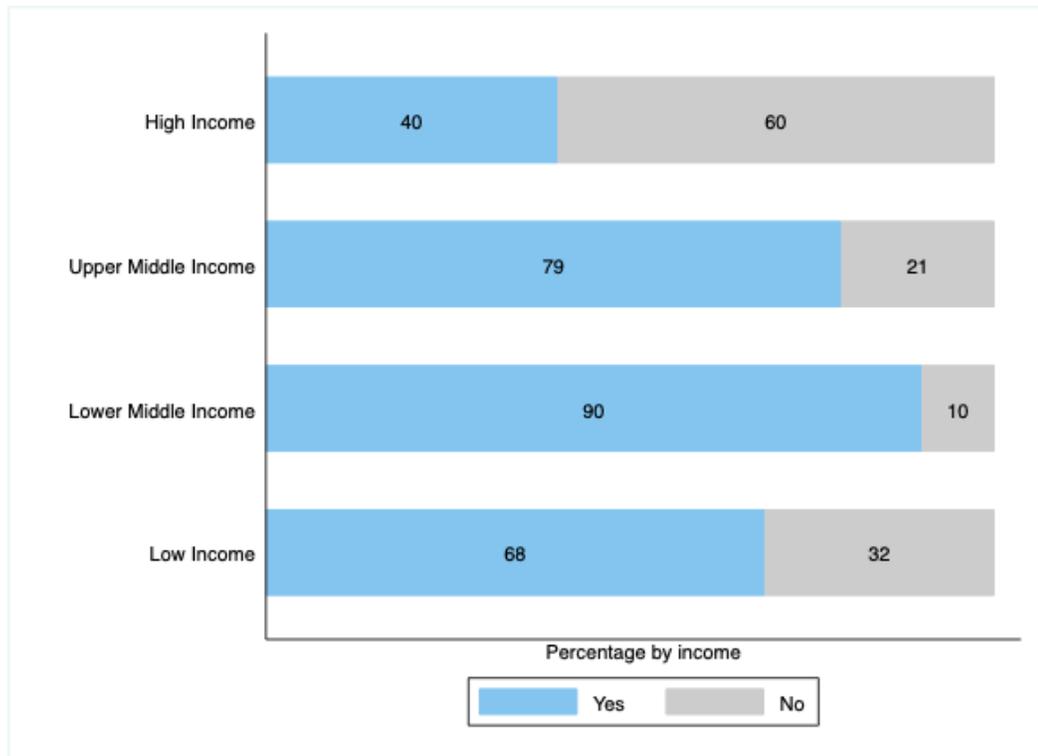


N = 76. External funding for a DA refers to receiving funding from donors, INGOs, private firms or other organizations outside national government for the establishment, design, and/or operationalization of the approach. This encompasses funding to set up structures and processes, salaries to compensate staff, funding or in-kind services for technical assistance for the design of the approach and consulting on best practices, and other financial support required to carry out the approach.

While we did not code the objectives of external funding or technical assistance, qualitative data from our cases indicates that external funding can be used for a range of purposes such as setting up of the DA structure, developing the terms of reference for the operations and staff responsibilities of the structure, and supporting staff recruitment or any technical assistance requirements. The European Commission, for example, provided both funding and technical assistance to the PMDU staff in Jordan in key skills related to monitoring and evaluation and problem-solving with line ministries (EuroAid, 2016; European Commission, 2018). In Sierra Leone, funding from the Global Partnership for Education (GPE) was used to set up a delivery unit to monitor progress towards the government’s priorities, assess implementation performance,

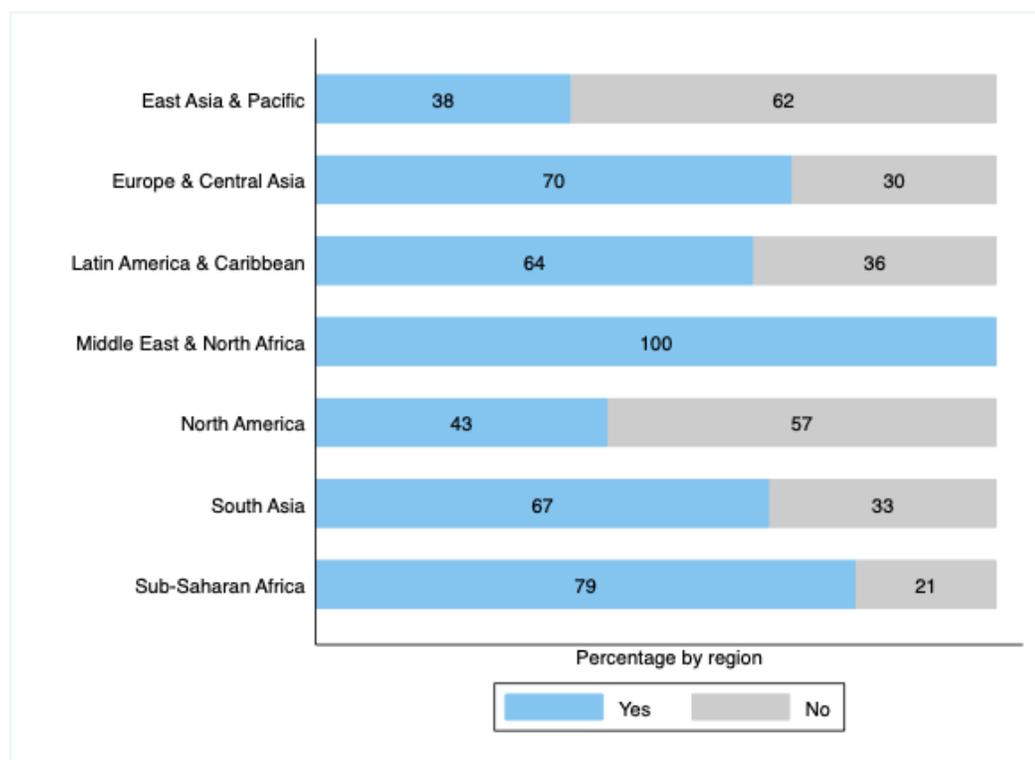
and preemptively resolve problems (GPE, 2020). In other cases, external funding is provided to supplement staff salaries or cover salaries of staff members entirely, especially if they are recruited as consultants (Todd, 2018). With respect to technical assistance, providing consultancy and technical advice to governments on delivery is a continually growing market. Since not all such relationships are disclosed publicly, our reported figures are likely an underestimate. Nonetheless, there is a clear pattern that most lower- and middle-income countries (and even many high-income countries) rely on some form of capacity building or advising in developing or operating their DA (see some of the key funding and consulting stakeholders in Gold’s 2017 report, *Tracking Delivery: Global Trends and Warning Signs in Delivery Units*).

Figure 11: DAS IN LOWER-INCOME REGIONS ARE MORE LIKELY TO RECEIVE TECHNICAL ASSISTANCE



N = 78. Technical assistance refers to the support that a DA receives to advise on the design and/or operations of the approach, the skills-training provided to staff working on the approach, or the human resources provided to supplement staffing of the DA.

Figure 12: IN MOST REGIONS, THE MAJORITY OF DAs RECEIVE TECHNICAL ASSISTANCE



$N = 78$. *Technical assistance* refers to the support that a DA receives to advise on the design and/or operations of the approach, the skills-training provided to staff working on the approach, or the human resources provided to supplement staffing of the DA.

5.2 What do delivery approaches aim to achieve and how?

This subsection discusses patterns in the functioning of DAs observed in our database, including the types of goals the DAs were set up to achieve and the time-horizon of these goals, and the mechanisms (i.e., managerial approaches) they sought to combine to try to improve bureaucratic functioning and policy implementation.

The main trends we observe are:

- The plurality of DAs (46%) focused mainly on achieving output-based goals (e.g., passing legislation, number of schools constructed), while 29% focused mainly on achieving

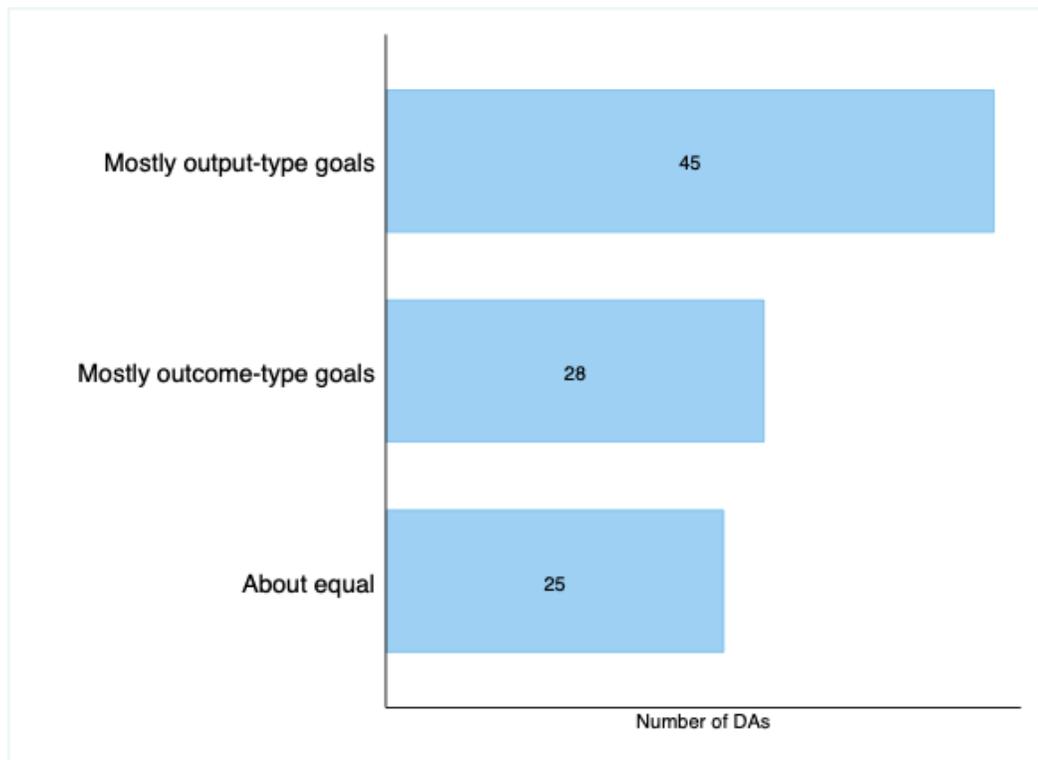
outcome-type goals (e.g., improving student learning, other citizen-focused metrics), and 25% focused more or less equally on both types of goals. DAs in middle- and low-income countries were more likely to define output-type goals than those in high-income countries.

- The majority of DAs (58%) relied more heavily on accountability- and incentive-driven mechanisms (e.g., through establishing top-down targets, incentives, and controls) to try to improve downstream bureaucratic performance, while 11% relied more heavily on problem-solving and organizational learning-driven mechanisms. Thirty-one (31) percent combined both pathways about equally. DAs in middle- and low-income countries were relatively more likely to utilize accountability- and incentive-driven mechanisms than those in high-income countries.

The remainder of this section defines these characteristics of DAs and presents descriptive statistics on worldwide trends, with selected illustrative examples.

A key distinction in public administration is between *outputs*, which are reforms, interventions, physical infrastructure, or actions produced by government bureaucracies themselves, and *outcomes*, which capture the end goals that government outputs aim to improve but which are not typically under a government's exclusive control (e.g., student learning outcomes, poverty rates). Examining the cases in our data, we find that almost half of DAs (46%) focused primarily on output-type goals such as implementing particular interventions or reforms. These goals include, for example, the number of hours of training delivered to teachers, number of schools constructed, number of emergency rooms opened in hospitals, or passing specific legislation. For example, the Results and Effectiveness (R&E) Unit embedded within the Royal Hashemite Court (RHC) in Jordan was charged with ensuring that the National Center for Curriculum Development (NCCD) was established within a year of the National Human Resource Strategy (HRD 2016-2025) being launched (NCHRD, 2016). Achieving this target entailed coordinating with several line ministries, including the Ministry of Education and the Ministry of Finance, in addition to ensuring that the necessary legislation was passed and the appropriate logistical arrangements made (Qarout, n.d.).

Figure 13: DAs MORE COMMONLY FOCUSED ON OUTPUTS-TYPE GOALS THAN OUTCOME-TYPE GOALS



Notes. N = 98. The goals monitored by the DA were coded based on whether they were mostly output-based (i.e., the set of materials, activities, or products produced over the course of service delivery, such as number of teacher training sessions conducted or number of schools and hospitals constructed), mostly outcome-based (i.e., the impact experienced by the beneficiary of the services provided, such as learning outcomes for students), or an approximately equal mix of both types.

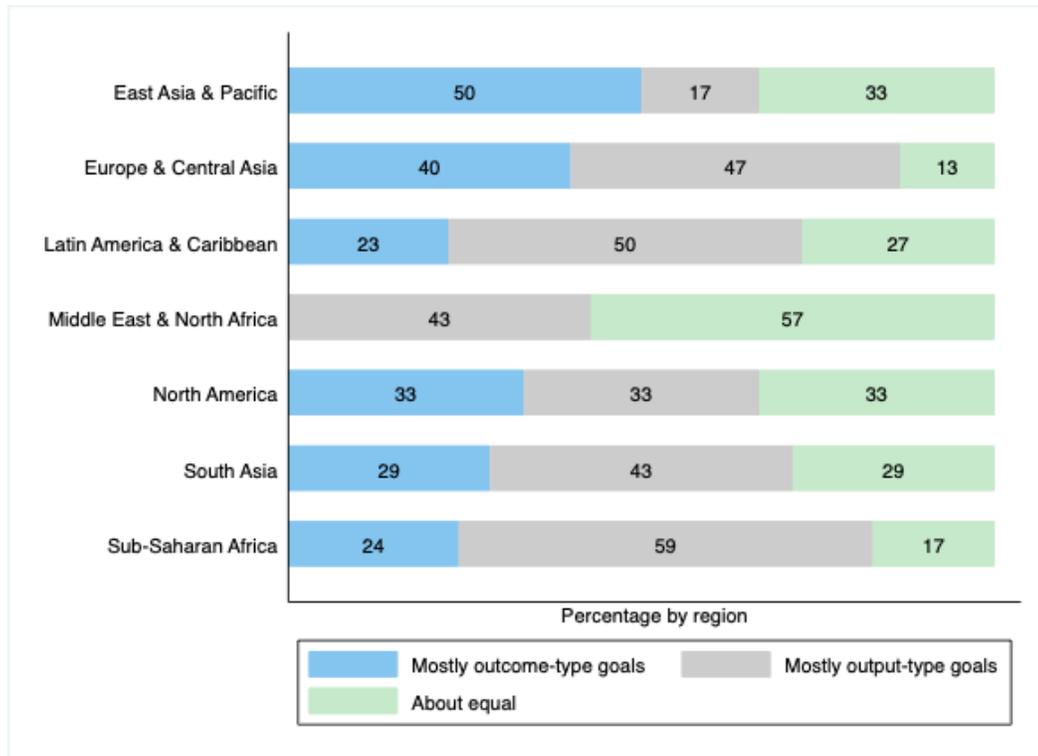
On the other hand, over a quarter of DAs (29%) focused mostly on outcome-type goals. For example, the DA in Pernambuco in Brazil monitored several outcomes-focused targets and was tasked with ensuring that some of the lowest performing schools in the state improved their rankings relative to other states in the national student examinations (Allesandro et al., 2014). The remaining 25% of cases combined both output- and outcome-type goals in approximately

equal measure.¹² Much of the existing literature on delivery units recommends focusing on goals that improve outcomes for service beneficiaries such as boosting literacy and numeracy among school children, increasing immunization rates, and increasing the number of new jobs in the labor market (e.g., Barber et al., 2011; Barber et al., 2011; Andrews, 2014; Barber, 2015), and other management literature has emphasized the importance of balancing goals across activity types and time horizons (e.g., Kaplan & Norton, 1992, 1993). While these prescriptions do not necessarily align with our output/outcome coding distinction, we note that the relatively high percentage of DAs focusing mainly on outputs (especially in lower-income countries, but also in high-income countries) suggests that either many governments are reluctant to foreground outcome-type goals in their DAs, or that some governments find DAs an attractive model for delivering on the output-type goals that are more standard within the public sector.

We observe variations in the output-vs-outcome goal balance across both regions (Figure 14) and income groupings (Figure 15). More than half (59%) of the DAs set up in Sub-Saharan Africa are mostly outputs-focused, higher than the rate in South Asia, North America, and East Asia and Pacific. On the other hand, half of the cases in East Asia and the Pacific focus on outcome-based goals, which is higher than all other regions, whereas no DAs in the MENA region focus mainly on outcome goals. More than three quarters of DAs in low-income countries (79%) monitor outputs-focused goals, which is a much higher portion of DAs than in other country income categories (see Figure 15), whereas higher-income countries are more likely to adopt mostly outcome-type goals. While this empirical pattern is broadly consistent with recommendations in the delivery unit and wider public administration literature that countries focus on the basics of reform implementation and output delivery before moving on to more sophisticated types of goals (e.g., Schick, 1998; Mourshed et al., 2010; Barber, 2015), the fact that we observe this pattern does not necessarily imply that this pattern of goal types is optimal. Whether DAs actually make these adaptations over time and the degree to which output vs. outcome design choices are beneficial/optimal in lower-income countries are important questions for future research, including through DeliverEd’s ongoing in-depth country research in Ghana, Jordan, Pakistan, and Tanzania.

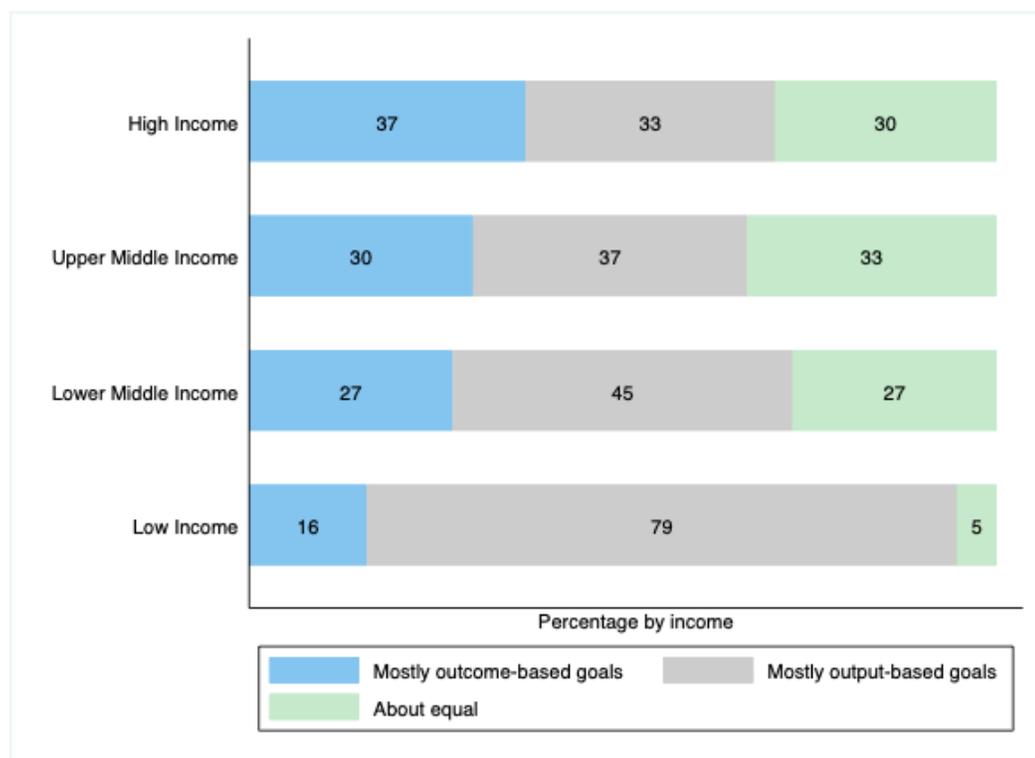
¹²Since not all DAs explicitly enumerated their goals and not all goals are equally important, we opted not to specify a numeric threshold for what constituted ‘mostly’ or ‘about equal’ with respect to goal type. We instead rely on coders’ judgment about the overall balance of goals espoused by each DA, with frequent discussion and harmonization among coders to ensure consistency.

Figure 14: DELIVERY APPROACHES VARY ACROSS COUNTRIES AND REGIONS IN THE TYPES OF GOALS ON WHICH THEY FOCUS



N = 98. The goals monitored by the DA were coded based on whether they were mostly output-based (i.e., the set of materials, activities, or products produced over the course of service delivery, such as number of teacher training sessions conducted or number of schools and hospitals constructed), mostly outcome-based (i.e., the impact experienced by the beneficiary of the services provided, such as learning outcomes for students), or an approximately equal mix of both types.

Figure 15: MORE THAN THREE QUARTERS OF DELIVERY APPROACHES IN LOW-INCOME COUNTRIES FOCUS ON OUTPUT-TYPE GOALS

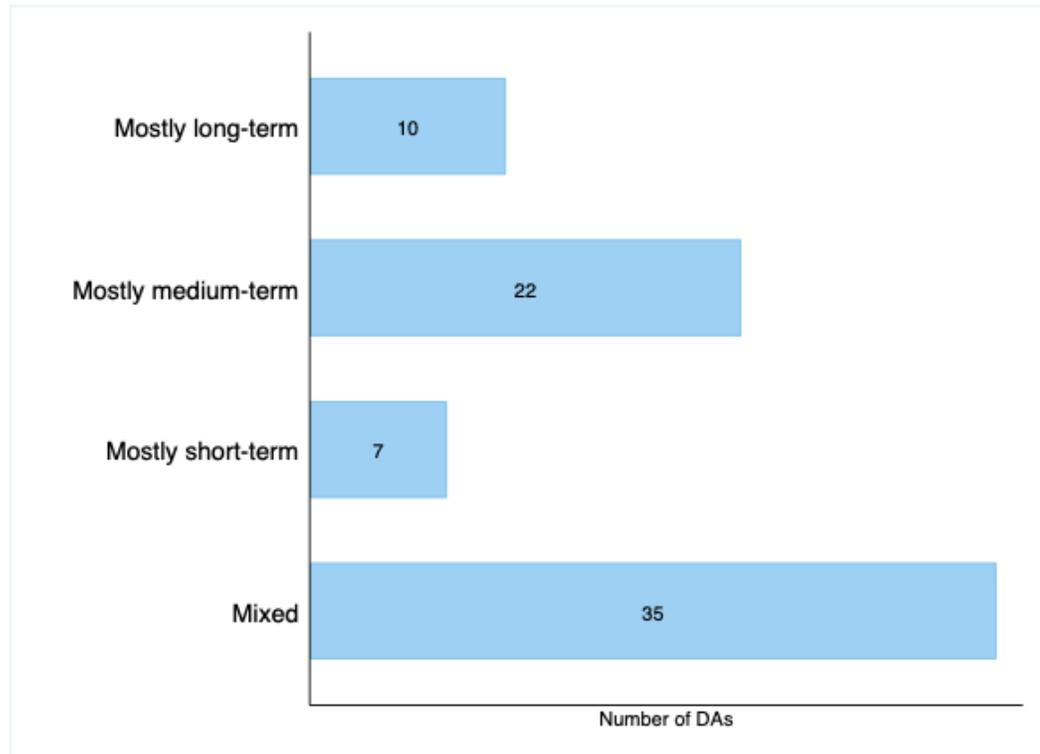


N = 98. The goals monitored by the DA were coded based on whether they were mostly output-based (i.e., the set of materials, activities, or products produced over the course of service delivery, such as number of teacher training sessions conducted or number of schools and hospitals constructed), mostly outcome-based (i.e., the impact experienced by the beneficiary of the services provided, such as learning outcomes for students), or an approximately equal mix of both types.

Figure 9 shows the time horizon of goals adopted by DAs and shows that almost half (47%) of the adopted goals appear to be of varying (mixed) time horizons. Of the DAs that focused primarily on a single time horizon, medium-term goals (3-5 years) were the most common, which coincides with typical electoral time horizons. A small percentage of DAs (9%) focuses solely on short-term goals (1-2 years), and only in a few cases (14%) did DAs focus primarily on longer time (6+ years) goals. This is broadly consistent with Gold’s (2017) observation that the general lifetime of a delivery unit is relatively short compared to other government structures, spanning from two to five years, and is often linked to the tenure of the political official who made the

decision to adopt the model.

Figure 16: DELIVERY APPROACHES TEND TO ADOPT GOALS WITH A MIX OF TIME HORIZONS



N = 74. The goal time horizon refers to the lifespan of the goals being monitored by the DA. Mostly short-term goals refer to goals that are intended to be implemented over 1-2 years. Mostly medium-term refers to goals meant to be implemented over 3-5 years and mostly long-term goals are intended to be implemented over 6 years or more. The coding was based on coder's estimation of the split between these three categories. If the split was more or less even, the case was coded 'mixed'.

A final design choice that governments adopting DAs need to make relates to its use of different managerial functions to try to improve policy implementation and bureaucratic performance. This choice will (or at least should) be guided by the government's understanding of the mechanisms that translate these functions into impact on policy delivery. Following an extensive literature in public administration, Williams et al. (2020) draw a conceptual distinction between mechanisms that emphasize creating routines related to accountability and incentives (e.g., through establishing top-down targets, incentives, and controls), and mechanisms that es-

establish routines for collaborative problem-solving and organizational learning.¹³ Based on this conceptual distinction, we use the secondary data identified through our global mapping process to investigate the extent to which DAs attempt to utilize and combine these mechanisms *in practice*.

To investigate patterns in the extent to which DAs worldwide attempt to make use of these two mechanisms, we use publicly available documents on each case to code whether each case uses ‘mostly accountability- and incentive-driven mechanisms’, ‘mostly problem solving- and organizational learning-driven mechanisms’, or draws on both ‘about equally’. This coding scheme recognizes that while most approaches do draw on both to at least some extent (as well as on other mechanisms which we do not attempt to code) and while there are potentially important interactions between these mechanisms (e.g., Andrews, 2014; Barber, 2015; Gold, 2017; World Bank, 2017; O’Malley, 2019), the overall balance between these two mechanisms can differ significantly across different DAs. While the overall patterns revealed at the global level are informative, we nonetheless caution that this relatively coarse coding scheme cannot capture all the intricacies and variations of these mechanisms’ design and interaction, that our coding (as with any large-scale coding exercise) is based on secondary rather than primary data, that these concepts exist on spectrums and that dividing them into discrete categories inevitably obscures some of this complexity, and that this exercise is purely descriptive and does not assume that any of these approaches is necessarily optimal. Finally, we note that ‘mostly’ and ‘about equal’ are admittedly imprecise terms, reflecting the fact that there is no natural way to measure these concepts; while they are a simple and intuitive way to analyze and communicate these global patterns, they are, nonetheless, limited in their explanatory power for any single case. Appendix B discusses this conceptual distinction and the coding process in more detail and gives illustrative examples.

Figure 17 shows that most DAs (58%) appear to rely relatively more heavily on accountability and incentive-driven mechanisms, that approaches focused primarily on problem solving and organizational learning are rare (11%), and 31% utilize both mechanisms in approximately equal

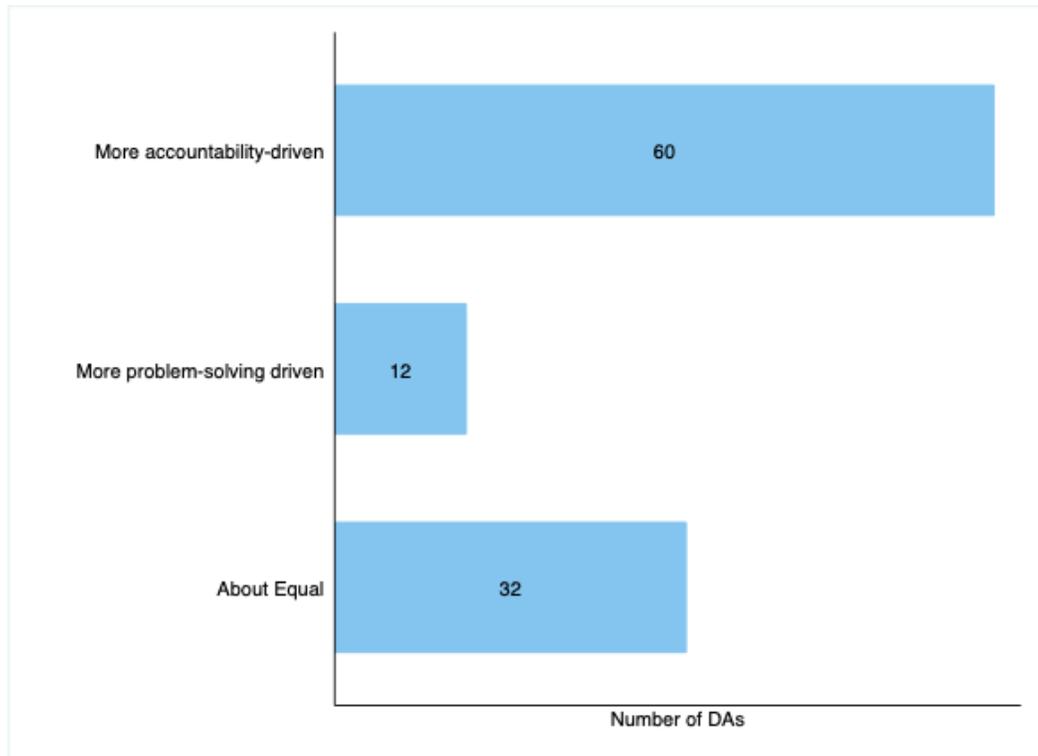
¹³Williams et al. (2020) refer to these two mechanisms as ‘Pathway A’ and ‘Pathway B’, respectively. The distinction, and potential trade-off between, these pathways has been a central feature of public administration scholarship for decades, with an extensive literature interrogating these issues through different methodological lenses and at different levels of government (see, e.g., Friedrich, 1940; Finer, 1941; Korten 1981; Simon 1983; Rose-Ackerman, 1986; Wilson, 1989; Rainey & Steinbauer, 1999; Carpenter, 2001; Gruening, 2001; Chun & Rainey, 2005; Duflo et al., 2012; Andersen & Moynihan., 2016; Andrews et al., 2017; Rasul & Rogger, 2018; Rasul et al., 2020). As Williams et al. (2020) emphasize, in practice these two approaches are not mutually exclusive and indeed are often combined (whether deliberately or not).

measure. Cases coded as adopting more accountability- and incentive-type approaches were those that tended to rely more heavily on monitoring routines that involved regular updates on agency performance with the political sponsor (president, prime minister, minister, mayor, or equivalent), tools that publicly signaled performance of various line ministries (e.g., using dashboards to track progress against targets), and various means (formal and informal) of creating rewards and/or sanctions for performance (firing or promotion, improving or risking reputation, etc.).

In comparison, approaches that were coded as relying more on problem-solving- and organizational learning-type approaches were those that engaged in frequent and timely troubleshooting of implementation challenges and that established lower-stakes collaborative forums that facilitated collective problem-solving and organizational learning. Existing reviews of delivery units that involve problem-solving routines point towards similar strategies as those highlighted above. For example, Gold (2017) discusses the use of deep-dive reviews which are organized to dissect implementation problems and plan appropriate responses. Furthermore, ‘communities of practice’ are also noted in the literature as one of the ways in which various stakeholders across agencies responsible for service delivery are able to convene and communicate key challenges for service delivery (e.g., Gold, 2017; McKay, 2017). At the same time, other authors have emphasized the potentially beneficial interactions between accountability and problem-solving and discussed the nuanced ways that both approaches can be integrated, as well as the potential challenges in doing so (e.g., Andrews, 2014; Barber, 2015; World Bank, 2017; O’Malley, 2019).¹⁴ These descriptive findings illustrate that different DAs do adopt substantively different approaches to using and combining these approaches, with further research needed to understand how best to do so in different contexts.

¹⁴See also Williams et al. (2020) for a discussion of the relationships among different managerial tools.

Figure 17: MOST DAs WORLDWIDE RELY MORE HEAVILY ON ACCOUNTABILITY AND INCENTIVE-DRIVEN MECHANISMS



N = 104. More accountability and incentive-driven DAs refer to cases where the main lever used in the approach are accountability mechanisms and routines that are intended to incentivize good performance (e.g., naming and shaming in public forums, regular updates with political sponsor to report on individual line ministry performance), while more problem-solving and organizational learning-driven approaches rely more heavily on routines for information sharing and problem-solving to enhance the implementation of reform (e.g., through iterative adaptation of delivery plans and timely information-sharing between front-line workers and central administration departments). These approaches are not mutually exclusive, although some DAs rely more on one type of approach than the other.

An example of a DA that used more problem-solving-driven mechanisms includes the U.S. State of Kentucky’s Commissioner’s Delivery Unit (CDU) embedded within the state’s department of education. The CDU relied heavily on data to set targets iteratively throughout the implementation process, identifying challenges in education service delivery, and developing solutions based on frontline needs (e.g., school administration, teachers, and students). The approach adopted

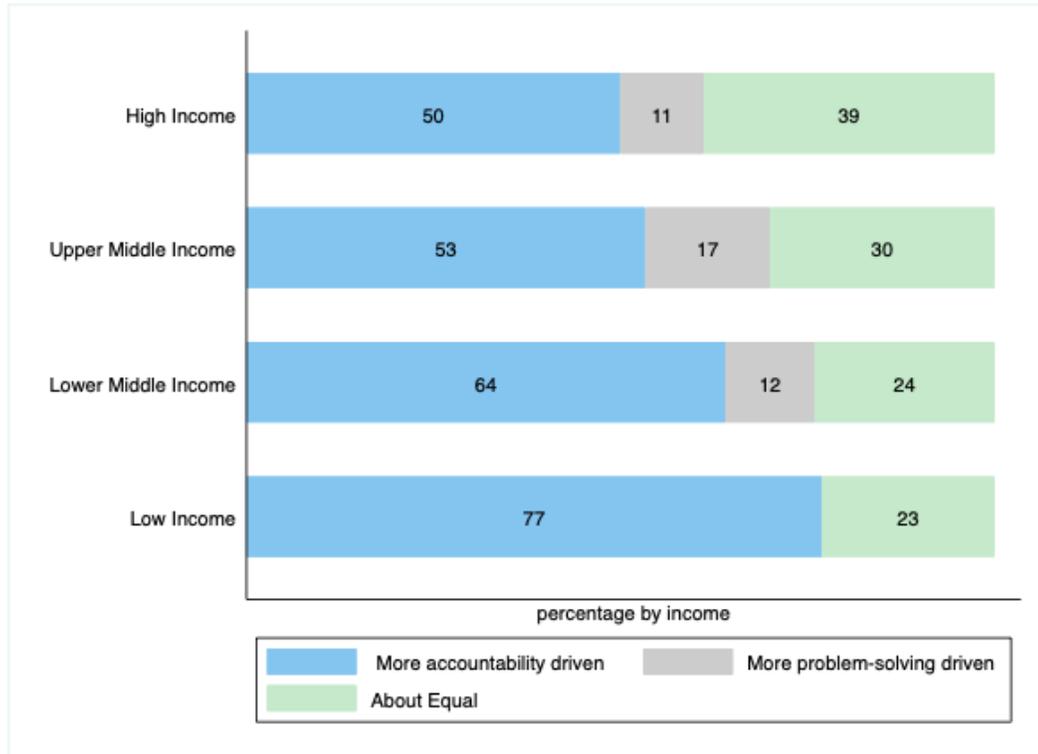
structured and systematic methods (such as Networked Improvement Communities (NICs) and Design-Based Implementation Research (DBIR), for example) to identify implementation challenges alongside peers, experts, and relevant stakeholders to inform revisions to the Department's delivery plans (McKay, 2017). For example, the unit sought to increase student preparedness for college by testing out opportunities for students to earn college credits in high school. The effectiveness of this program stemmed from the CDU's continuous study of student participation in the various dual enrollment programs schools offered and the performance demonstrated by enrolled students (McKay, 2017). The CDU relied heavily on the information shared by school administrations and a series of consultation sessions to identify progress made against the college preparation target the Department set for them (McKay, 2017).

DAs that exhibit a combination of both types of managerial approaches (i.e., accountability/incentives and problem-solving/organizational learning) leverage top-down incentives and accountability for results, as well as vertical and horizontal information-sharing and problem-solving across the delivery chain and with citizens. An example of a DA that employs both types of managerial approaches is Albania's Delivery Unit. The unit carried out routine stock-take meetings on a monthly basis with the Prime Minister and line Ministers along with maintaining the Delivery Dashboard that measured progress of implementation against pre-set indicators (Hart, 2017). In addition to these top-down monitoring and accountability management tools, the unit organized Delivery Challenge Sessions with key stakeholders along the delivery chain involved in achieving sectoral goals (Hart, 2017). These sessions were facilitated by the unit and produced Delivery Chain Assessments which identified the gaps and opportunities in existing delivery efforts. Malaysia's PEMANDU is another well-known example that combines elements of both mechanisms, using a progress dashboard to keep political sponsors informed of bottlenecks and successes of implementation, as well as deploying policy labs to engage beneficiaries of public services in setting national priorities and demanding specific policies and programs to meet their needs (World Bank, 2017; PEMANDU, n.d.).

Figures 18 and 19 show the adoption of managerial approach by region and country income level respectively. Low-income countries are more likely to adopt more accountability-intensive DAs compared to countries in other income categories, and none of the cases identified in this income category were coded as utilizing more problem-solving-intensive DAs. On the other hand, upper middle- and high-income countries are more likely to establish DAs that attempt to use both mechanisms to approximately equal extents or that focus primarily on problem-solving and organizational learning. These patterns are broadly replicated across regional groupings, albeit

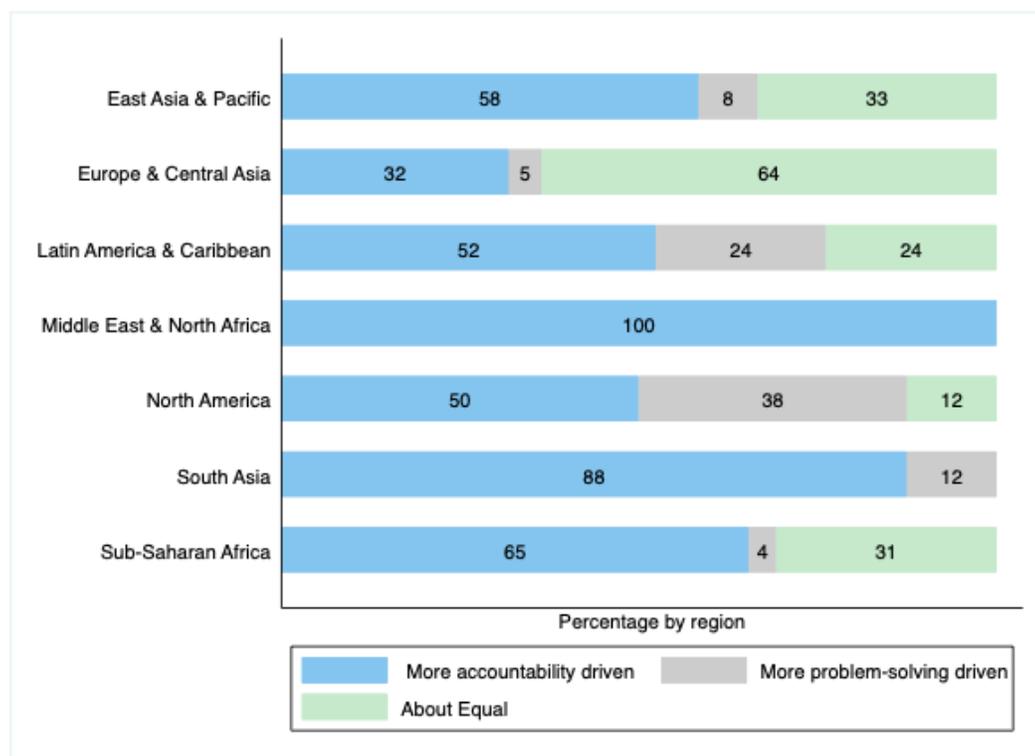
with certain regions (MENA and South Asia) perhaps more dominated by primarily incentives-driven approaches than their income levels might predict.

Figure 18: DELIVERY APPROACHES IN LOW-INCOME COUNTRIES ARE MORE LIKELY TO EMPHASIZE ACCOUNTABILITY- AND INCENTIVE-DRIVEN MECHANISMS



N = 104. More accountability- and incentive-driven delivery approaches refer to cases where the main lever used in the approach are accountability mechanisms and routines that are intended to incentivize good performance (e.g., naming and shaming in public forums, regular updates with political sponsor to report on individual line ministry performance), while more problem-solving and organizational learning-driven approaches rely more heavily on routines for information sharing and problem solving to enhance the implementation of reform (e.g., through iterative adaptation of delivery plans and timely information-sharing between frontline workers and central administration departments). These approaches are not mutually exclusive, although some DAs rely more on one type of approach than the other.

Figure 19: DELIVERY APPROACHES IN SOUTH ASIA AND THE MENA REGION ARE MOST LIKELY TO RELY MAINLY ON ACCOUNTABILITY- AND INCENTIVE-DRIVEN MECHANISMS

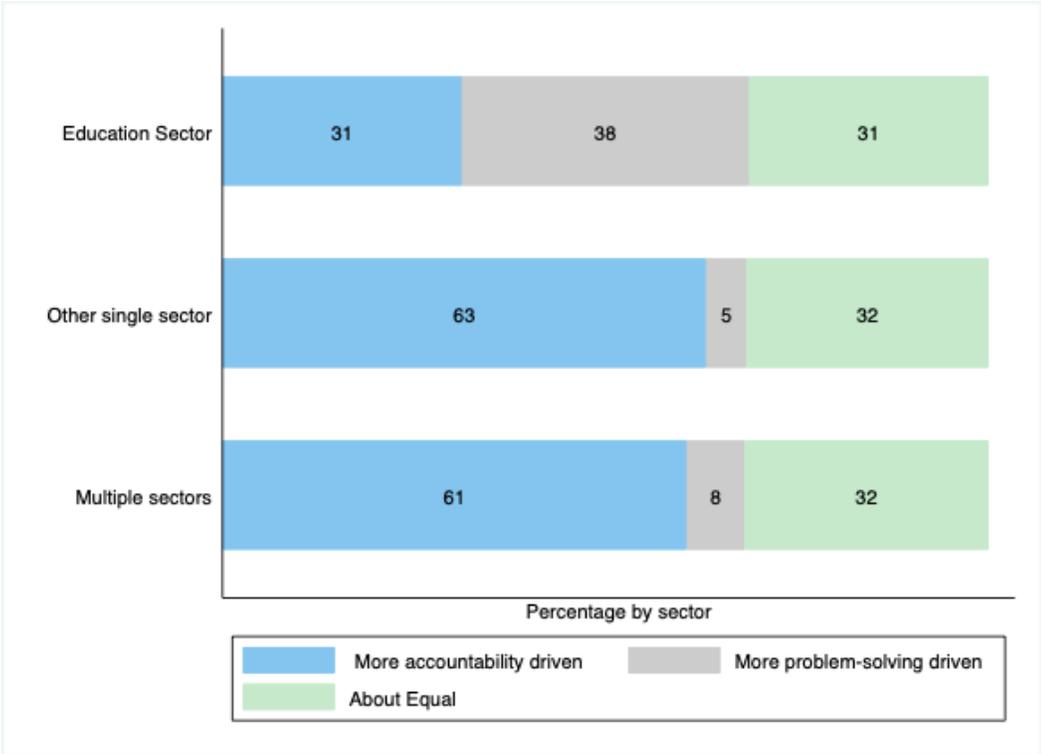


$N = 104$. More accountability- and incentive-driven delivery approaches refer to cases where the main lever used in the approach are accountability mechanisms and routines that are intended to incentivize good performance (e.g., naming and shaming in public forums, regular updates with political sponsor to report on individual line ministry performance), while more problem-solving and organizational learning-driven approaches rely more heavily on routines for information sharing and problem solving to enhance the implementation of reform (e.g., through iterative adaptation of delivery plans and timely information-sharing between frontline workers and central administration departments). These approaches are not mutually exclusive, although some DAs rely more on one type of approach than the other.

DAs that operate across multiple sectors (as opposed to within single sectors) are more likely to adopt more accountability- and incentive-heavy approaches to improve implementation, as Figure 20 shows. The exception appears for DAs adopted within the education sector, which has an unusually high concentration of DAs that focus mostly on problem-solving-driven approaches. Since DAs in other single sector were not systematically coded in the same way as

education-focused DAs, we cannot make conclusive claims about management norms that are more prominent in the education sector compared to others. Drawing on findings from other research on delivery units (e.g., Kohli et al., 2016; Gold, 2017; Barber, 2015; O'Malley, 2019), we can note that with a wider remit assigned to DAs, sustaining attention and deeper engagement with various segments of the delivery chain becomes more challenging. For example, the PMDU in Jordan suffered from stretching their resources and ambitions too thin across a wide range of government reform initiatives covering all 120 public sector agencies. What was intended to be a problem-solving-intensive DA transformed, in practice, into a monitoring and tracking check-list for staff at the PMO who collected data from line ministries on specific key indicators on a monthly basis (Qarout, n.d.; NCHRD, 2016). The main challenges the PMDU faced was allocating sufficient human resources to regularly engage with civil servants across ministries and to analyze data collected by various ministries on implementation progress (Qarout, n.d.).

Figure 20: DELIVERY APPROACHES WITHIN THE EDUCATION SECTOR TEND TO ADOPT MORE PROBLEM-SOLVING MANAGERIAL MECHANISMS COMPARED TO MULTISECTOR DAs



N = 101. More accountability- and incentive-driven delivery approaches refer to cases where the main lever used in the approach are accountability mechanisms and routines that are intended to incentivize good performance (e.g., naming and shaming in public forums, regular updates with political sponsor to report on individual line ministry performance), while more problem-solving and organizational learning-driven approaches rely more heavily on routines for information sharing and problem-solving to enhance the implementation of reform (e.g., through iterative adaptation of delivery plans and timely information-sharing between frontline workers and central administration departments). These approaches are not mutually exclusive, although some DAs rely more on one type of approach than the other.

6 Relationships Between Design Features

Our descriptive analysis in Sections 4 and 5 has illustrated the global patterns observed across a range of characteristics capturing the goals that are adopted by DAs, how DAs are located within government structures and processes, and the different mechanisms through which DAs aim to achieve their impact. In this section, we investigate the relationships among these design features, to explore, for instance, whether governments tend to choose different types of managerial approaches to achieve different types of goals. To do so, we explore the empirical correlations among selected features of DAs in our dataset. We document the following relationships:

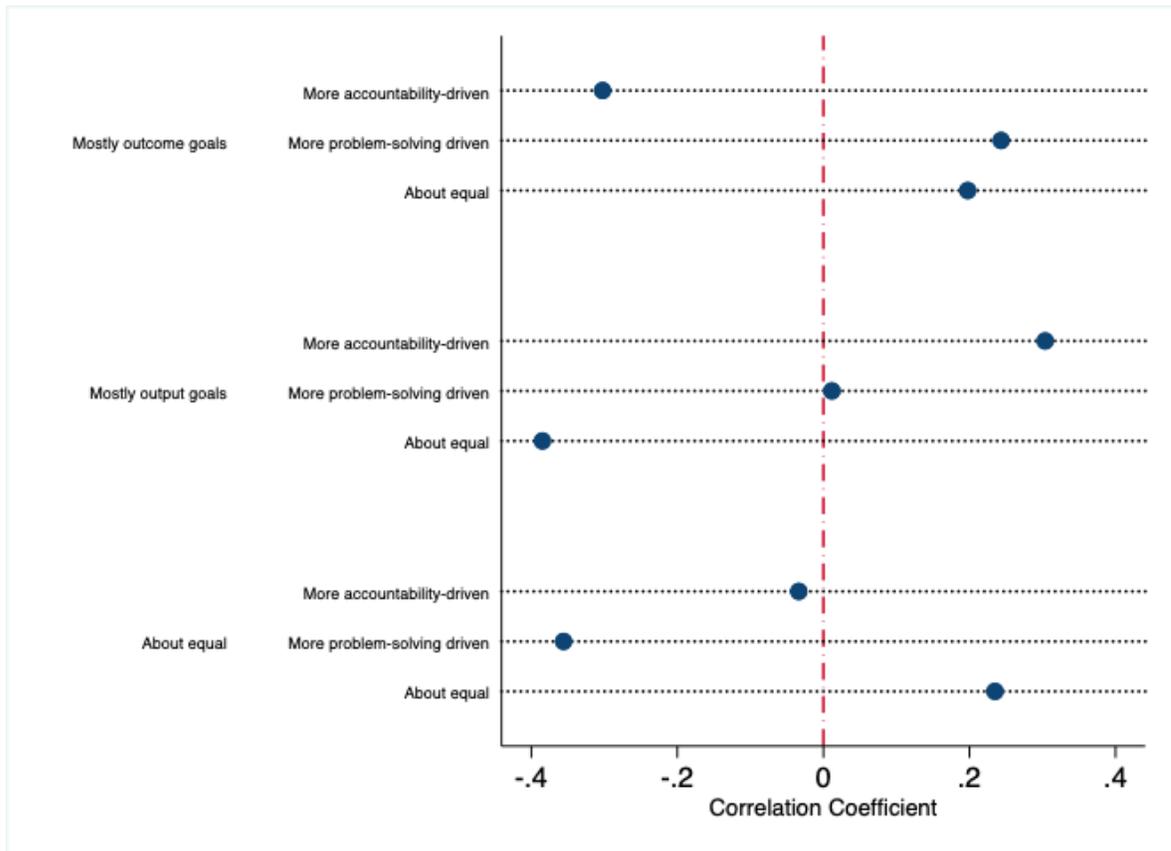
- DAs that rely more heavily on managerial approaches that involve accountability and incentives as levers for improvement tend to focus more on delivering specific outputs or reforms. In comparison, DAs that rely more heavily on managerial approaches that involve problem-solving and organizational learning or a balanced combination of these two managerial approaches tend to focus more on achieving improvements in outcomes or an equal mix of outcomes and outputs (rather than outputs or predefined reforms).
- DAs that rely more on accountability- and incentives-type managerial approaches are more likely to be housed within preexisting government structures, whereas DAs that employ more problem-solving- and organizational learning-type managerial approaches (or an equal mix of both accountability- and problem-solving-type routines) tend to exist in units that have been newly established or reorganized.
- DAs that are housed within preexisting government structures also tend to recruit staff exclusively from within the civil service. On the other hand, DAs that are housed in either new or reorganized structures tend to recruit entirely from outside the civil service or have mixed staff (i.e., a combination of civil servants and recruits from outside the civil service).

While these relationships provide useful insight into how various governments have combined various potential design choices for DAs, we emphasize the need for caution in interpreting them. First, the fact that these correlations between design features exist does not mean that they are necessarily optimal, either in general or in specific cases. Second, we present simple bivariate correlations, leaving potential investigation of multivariate relationships for future work. Third, the observed correlations do not necessarily imply causal relationships between variables, both

because of the potential for omitted variable bias and because many of these design choices are simultaneously determined. Finally, the fact that the number of DAs worldwide documented in our database are limited means that not all of these correlations are statistically significant at conventional levels, but we present them because they are based on the universe of DAs identified and document what appear to be important empirical trends. In the remainder of this section, we discuss each of these relationships in detail.

With respect to the relationship between DA managerial approach and goal types, Figure 21 shows that DAs with managerial approaches that rely more heavily on accountability- and incentive-driven routines tend to be those that focus more on delivering specific outputs or reforms, whereas DAs that rely more heavily on problem-solving and organizational learning or an equal mix of both types of managerial approaches tend to be those that focus more on achieving outcomes or on a mix of output and outcome-related goals. This relates to existing empirical literature on performance-based incentives which highlights that these are most effective when they can be linked to clear and explicit measures (e.g., Bandiera et al., 2007; Khan et al., 2016) and the public administration literature that links managerial strategies to goal type (e.g. Wilson, 1989; Chun & Rainey, 2005; Rasul & Rogger, 2018; Rasul et al., 2020). While specific output-related goals (such as number of textbooks or trainings delivered) can be measured explicitly, allowing for effective incentive schemes to be designed, outcome-related goals (such as quality of teaching for example), in comparison, are often more challenging to measure which can make designing and administering incentives harder. This can explain why managerial approaches that focus more on accountability and incentive routines tend to co-exist more commonly with output-focused goals.

Figure 21: CORRELATIONS BETWEEN DA MANAGERIAL APPROACH AND GOAL TYPES

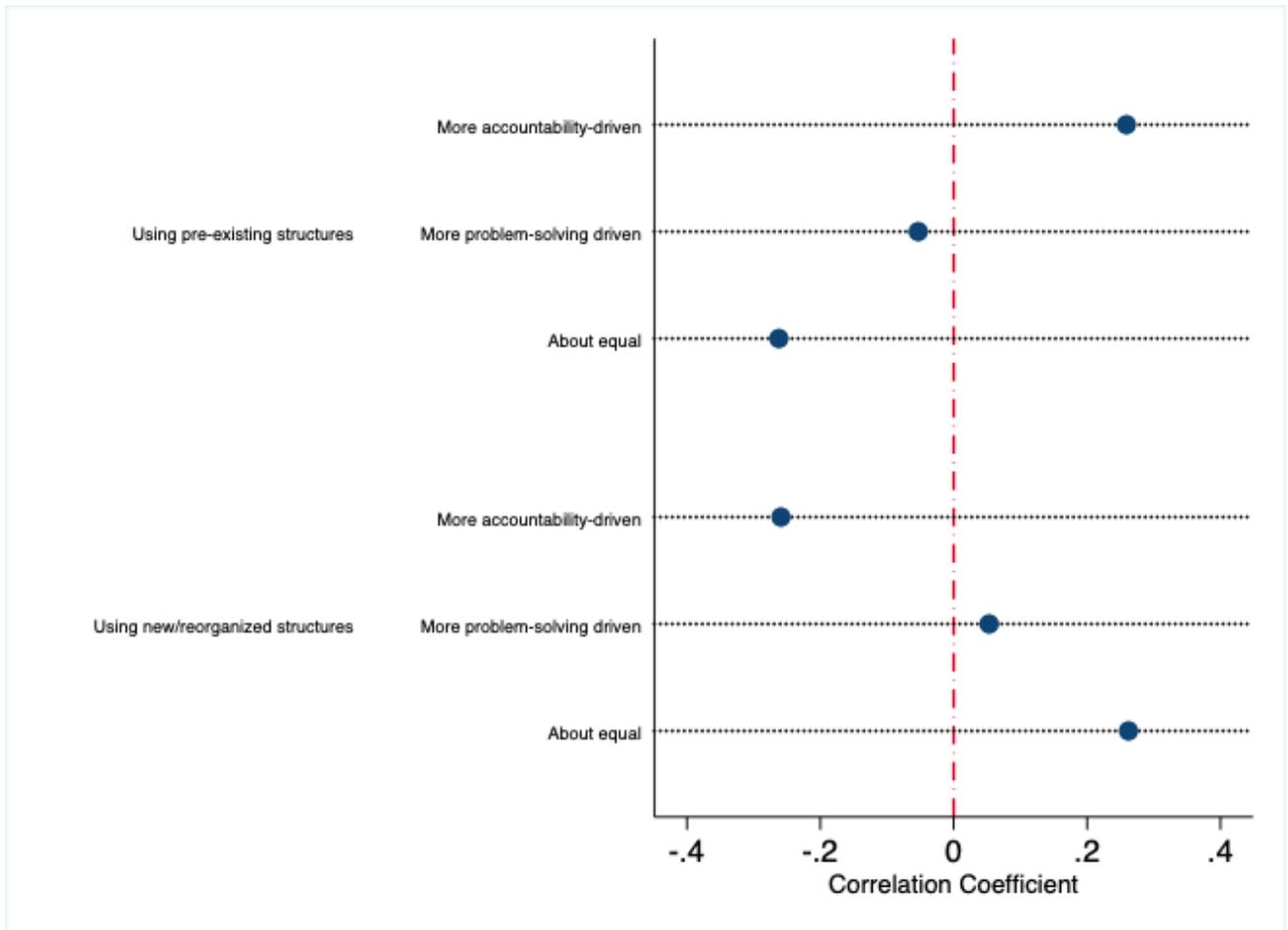


Notes. $N=85$. The figure displays 9 different correlation coefficients between each type of DA managerial approach (more accountability- and incentive- driven, more problem-solving driven, about equal) and each type of DA goal (mostly outcome, mostly output, about equal). Each design feature and its type is coded using dummy variables.

One example of a DA that illustrates this broader pattern is the Implementation Support and Follow Up Unit (ISFU) in Oman (2016-present), which relies mostly on accountability- and incentive-driven managerial approaches. These align with the goals that the ISFU monitors, which are mainly outputs-focused and are relatively easier to measure such as completion of the master plan of the Khazaen Economic City, completion of framework for domestic tourism, and administration of training for legal researchers at the Ministry of Manpower, to name a few (ISFU, 2017).

The type of managerial approach adopted by DAs is also correlated with the choice of whether to house the DA in an existing part of the bureaucracy rather than creating a new or reorganized unit. Figure 22 shows that DAs that are housed in preexisting structures are more likely to rely more heavily on accountability- and incentives-driven managerial approaches. On the other hand, DAs that rely more heavily on problem-solving and organizational learning-driven managerial approaches or an equal mix of these two types of approaches tend to be housed in new or reorganized structures (e.g. merging of different departments or a reorganisation of the department's organizational chart). Since DAs that employ problem-solving driven approaches (either exclusively or in combination with accountability driven approaches) often require new platforms or new ways of thinking to understand and resolve implementation challenges, it is possible that new or reorganised structures make it relatively easier to adopt such innovative approaches. Other potential alternative explanations for this relationship are that mainstream bureaucratic structures are better able to leverage accountability relationships since these typically already exist, or that DAs housed within preexisting structures are more likely to be established when delivery priorities are output-type in nature and thus more likely to adopt accountability-focused approaches.

Figure 22: CORRELATIONS BETWEEN DA MANAGERIAL APPROACH AND STRUCTURE



Notes. $N = 92$. The figure displays 6 different correlation coefficients between each type of DA managerial approach (more accountability and incentive heavy, more problem-solving heavy, about equal) and each type of DA structure (housed in preexisting structure, new or re-organised structure). Each design feature and its type is coded using dummy variables.

An example that illustrates this trend is the Cape Town Strategy and Policy Unit (SPU), which was set up in 2011. Before it was established, the city government had been trying to address the thorny challenge of how to facilitate deeper collaboration with the private sector as part of the city’s social development strategy. According to the director of the unit, previously, staff were “unsure about where the boundaries would be and the mindset in general was one of anxiety about this type of innovation” (Center for Public Impact, 2016). After the new Cape

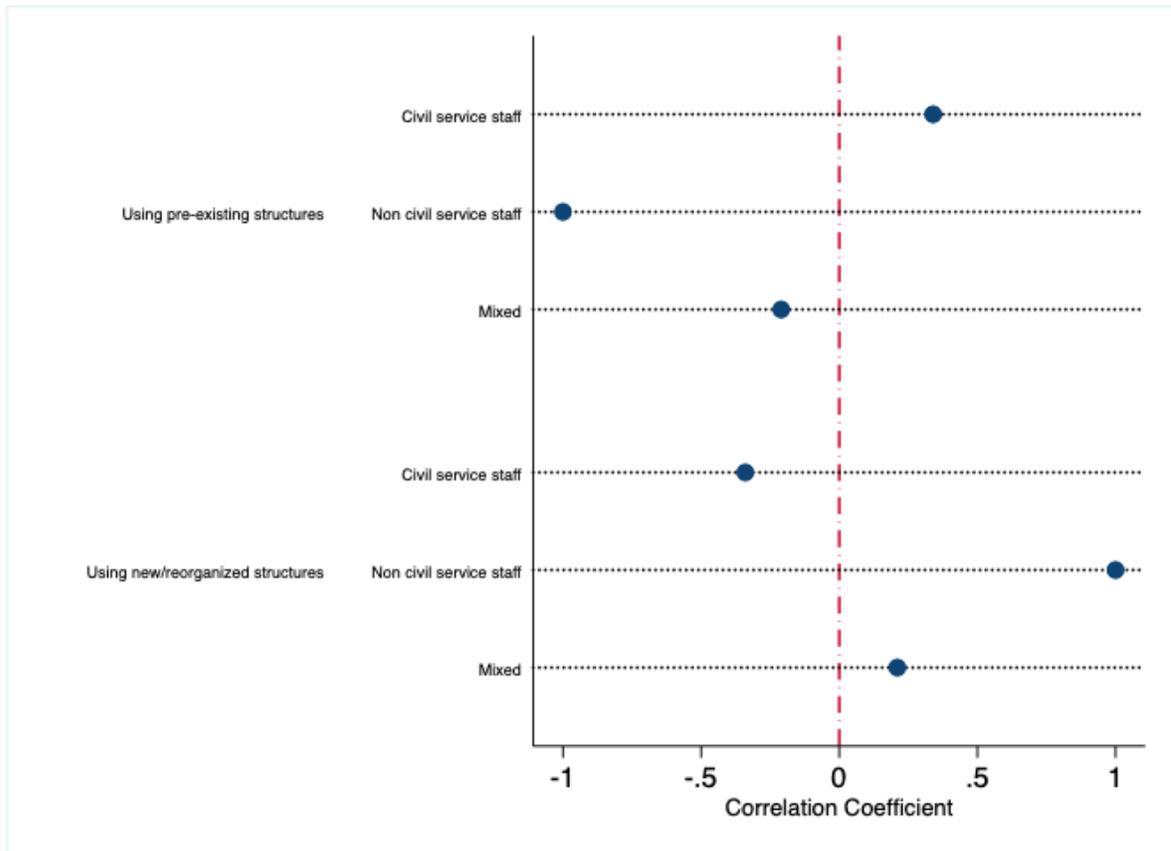
Town mayor reorganized its existing bureaucracy to set up a new Strategic Policy Unit (SPU), staff were better able to cut across preexisting departmental boundaries and engage with actors from different local governments departments (such as legal and supply chain management) to understand how to break communication barriers (Center for Public Impact, 2016). The new ambit of the SPU appeared to make it relatively easier to adopt a fresh mandate and new ways of working.

Perhaps unsurprisingly, Figure 23 shows that DAs that rely more on preexisting bureaucratic structures also tend to be staffed from within the civil service, whereas DAs that are set up as new or reorganised structures are more likely to have non civil service or mixed staff (i.e., a combination of civil service and non-civil service staff). The latter relationship is also illustrated by insights from policymakers involved in the setting up of the Strategy and Policy Unit (SPU) in Rwanda. They point out that the significant restructuring of its center of government to set up the unit had to go hand in hand with a simultaneous increase in individual skill sets for planning, management, and a range of other capabilities of staff (Iyer, 2012), which was made possible through a mixed staffing structure at the SPU. Relatedly, insights from other authors such as Barber (2015) that underline the need to recruit from the private sector to supplement government efforts in carrying out delivery reforms are also relevant for understanding these relationships. Given government departments often have rigidly defined staffing structures, such forms of external hiring could practically be more feasible in DAs that are either newly set-up or reorganized.

While external staff can help bring new skill sets to the table and expedite short-term action, they can also pose challenges of institutional resistance and buy-in from core civil service staff. For example, the Strategy and Policy Unit (SPU) in Sierra Leone (2008-2012) was initially set up through hiring external advisors and analysts. The lack of representation of the civil service, along with the absence of a future-looking plan of how the unit would eventually incorporate the civil service into its structure, created resistance to the unit by other line ministries and stakeholders (ACET 2010).¹⁵ While any DA that works with several line ministries and/or across subnational management structures (e.g., in federal systems) possibly has to invest time in building its own credibility and earn the trust of its stakeholders, this is likely to be even more critical for DAs that operate outside of normal civil service structures and/or staffing procedures.

¹⁵An evaluation published by the African Development Bank Group later in 2013 suggests that based on the recommendations in ACET (2010), civil servants were offered the opportunity to be seconded to the SPU to mitigate the tension between SPU staff and civil servants.

Figure 23: CORRELATIONS BETWEEN DA STRUCTURE AND STAFFING



Notes. $N = 66$. The figure displays 6 different correlation coefficients between each type of DA structure (using preexisting bureaucratic structures, using new or reorganized bureaucratic structures) and each type of DA staffing (civil service staff, non-civil service staff, mixed). Each design feature and its type is coded using dummy variables.

While the correlations presented above highlight common relationships between different design features in our dataset, two points are pertinent to mention. First, there are in fact many types of DAs that do not conform to these patterns, such as DAs that employ more accountability- and incentive-driven approaches but focus mostly on outcome-focused goals. Therefore, the correlations presented above document that some design choices of policymakers tend to go hand-in-hand, but these relationships are by no means universal. Second, and relatedly, these correlations do not necessarily establish the optimal way of making design choices together. While some of these can be related to existing literature or appear practically intuitive, whether

or not these combination of choices are in fact the most efficient or optimal remains a question for further empirical inquiry.

7 Alternative Delivery Approaches Designs

We have focused our analysis in Sections 4-6 on DAs that consist of a specific institutionalized unit (for instance delivery units or implementation units). However, our definition of a DA (see Section 2) also emphasises that DAs can also consist of “structured processes” rather than specific units with their own organizational form. Although structured process DAs share many of the goals and managerial functions of institutionalized unit DAs, we excluded such cases from our main analysis since they were more difficult to identify through our systematic mapping procedure and thus comprised just 6% of all DAs documented worldwide. Structured process DAs also represent a highly diverse category, making it even more difficult to identify patterns and trends than for institutionalised unit DAs.

In this section, we present qualitative descriptions of three examples of such structured process DAs to further illustrate the breadth of the options facing governments that are considering adopting a DA. The structured process DAs we observe mostly take the form of committees, commissions, new task forces, or a set of activities that bring together various stakeholders to carry out the role of setting up and implementing a DA. These cases were found in the following jurisdictions: Bahrain, Kuwait, Morocco, Thailand, Botswana, India, Cameroon, Mexico City, Denmark, and the United Arab Emirates.

In Kuwait, the government set up the National Sustainable Development Committee (NSDC) to monitor the implementation of the Vision 2035 and the country’s progress towards the Sustainable Development Goals (SDGs) (KVNR, 2019). Process-type approaches are often leveraged to coordinate across a large number of stakeholders to build consensus among various actors sharing the responsibility - and accountability - for implementation of reform. For example, the NSDC was established after the General Secretariat Supreme Council for Planning and Development (GSSCPD) and the Central Statistical Bureau (CSB) ran more than 13 nationwide workshops with over 66 stakeholders from government agencies, Non-Governmental Organizations, and private firms involved in the implementation of projects linked to SDGs. These workshops were organized to identify key national priorities and implementation plans with distinct roles and

responsibilities for project implementation across stakeholders (KVNR, 2019). The NSDC is, therefore, intended to ensure effective participation of representative stakeholders from public, private, and development organizations in the planning and delivery process. Civil society and line ministries are represented on the Committee, as well as international development and donor institutions such as the UN and the World Bank.

The Linking Education and Accountability for Development in Morocco is another example of an innovative DA that operates as a structured process. What started off as a project funded by the World Bank turned into an institutionalized process that the government owns and uses to monitor allocation of resources to schools, student-teacher assignments, teacher attendance, school maintenance, and procurement of materials for schools (World Bank, 2014). The DA adopts a participatory approach to school improvement by developing mechanisms through which parents are able to collaborate with school, district, regional, and central administrations within government (World Bank, 2014). Hence, it operates much more directly with frontline service deliverers than most DAs, which tend to focus on improving policy implementation and bureaucratic functioning farther up the delivery chain.

A third type of alternative approach is represented by the innovation labs operated in Mexico City and Denmark, which seek to mobilize various actors across government to identify problems and collectively brainstorm solutions to improve implementation (Buerkli, 2015; LABCD, n.d.). While these consultative and information-sharing activities are often organised by the CoG or another apex authority, they tend to rely significantly less than unit-style DAs on leveraging a single central team or a permanent physical structure that is responsible for ensuring implementation of policy and programs. While the innovation lab in Mexico City relies heavily on passing knowledge and feedback upward from downstream implementing partners, stakeholders from civil society, and citizens, the Mindlab in Denmark was a centrally driven initiative that coordinated across line ministries and facilitated expert knowledge-sharing and problem-solving among bureaucrats (Buerkli, 2015). These examples illustrate that the conceptual distinction between institutionalized unit and structured process DAs is blurry in practice, as many DAs combine some form of central unit or hub with reliance on multistakeholder processes.

These examples illustrate several points. First, DAs - efforts by government to use novel institutional forms to improve policy delivery by using a range of managerial tools to shift focus from inputs and processes to outputs and outcomes - can take a diverse range of forms, among which establishing a dedicated delivery unit in the core of government is only one. This means that

governments that want to adopt a DA have numerous options to choose from when deciding how to design them, even beyond the spectrum of variables discussed in Section 5. Second, the boundaries between DAs and other public administration tools or ideas, such as participatory policy-making, innovation promotion, and strategic policy-making, are inevitably blurred. While maintaining a conceptual distinction between a DA and other types of efforts to improve service delivery is important, DAs, in practice, are a diverse set of approaches that exist alongside and overlap with other public administration tools. Given this wide array of possibilities for policy-makers choosing how to configure the structure and functions of a DA, a key topic for future research is the relative effectiveness of different models in different contexts.

8 Conclusion

This paper has documented patterns in the adoption and design of delivery approaches (DA) aimed at improving service delivery worldwide over the past 25 years. We defined a DA as “an institutionalized unit or structured process within a government bureaucracy that aims to rapidly improve bureaucratic functioning and policy delivery by combining a set of managerial functions in a novel way to shift attention from inputs and processes to outputs and outcomes” (Williams et al., 2020). To document these patterns, we conducted a systematic worldwide mapping of cases of DA adoption at national government levels, supplemented by examples from selected subnational jurisdictions. Based on publicly available secondary data, we coded key features of these cases and examined qualitative data on each case. While our analysis has important methodological limitations and we are cautious about using it to infer lessons for whether and how governments ought to use delivery approaches, it nonetheless provides a strong basis on which to identify patterns in when, where, and how delivery approaches have been used around the world.

Our key finding is that while the adoption of DAs is widespread across countries and regions and has accelerated in the last decade, there is no single model of a delivery approach. Cases vary across numerous variables related to their design, such as the type and time horizon of their goals, their staffing and structure, and the main managerial levers they use to effect change. Furthermore, while we document empirical correlations between some of these features of delivery approaches and their contexts (for example, how the predominant managerial levers used by the DAs tend to vary across goal type and the approach’s structural integration into the preexisting

bureaucracy), there are nearly as many cases that do not conform to these patterns. While we do not attempt a systematic exploration of what factors shape governments' choices in adopting and designing DAs nor do we make claims about the relative effectiveness of different approaches, the sheer diversity of forms taken by DAs is evidence of the idiosyncratic nature of many of these decisions.

For policymakers interested in adopting a DA, there is thus a vast menu of potential options to choose from, both in terms of each design choice that has to be made, but also in the various ways in which these choices can be combined. This should encourage policymakers to be creative, as well as attentive to their specific goals and context, rather than simply copying models that appear to have been successful elsewhere. At the same time, this range of choices emphasizes the complexity of designing DAs. And while there is almost certainly no single optimal way in which to design a DA, the key question for policymakers and for future research alike is whether and why DAs work to improve service delivery, and how their design interacts with their intended goals and key features of the context.

Although this paper is descriptive in nature, research from the country-level component of the DeliverEd Initiative is investigating such questions through rigorous case studies across Tanzania, Pakistan, Ghana, and Jordan. The descriptive patterns of DAs documented through this paper highlight the variation and nuances in different design features of DAs that the country-level studies can take into account while designing and analyzing their research hypotheses and findings. This country-level research will then generate causal evidence on how selected features of DAs and their contextual environment are related to changes in bureaucratic functioning, policy implementation, and service delivery, which can be contextualized within the trends highlighted in this paper's broad mapping of the landscape of DAs worldwide. Together these research pieces will offer governments and policymakers insights into the menu of possible design options for setting up DAs, and rigorous evidence about their effectiveness and which types of design features are most and least suited to different contexts and goals.

References

- ACET (African Center for Economic Transformation). (2010). *Functional review of the Strategy and Policy Unit, Sierra Leone*. ACET.
- Allesandro, M., Lafuente, M., & Santiso, C. (2014). *The role of the center of government: A literature review*. Inter-American Development Bank.
<https://bit.ly/3hW70Rz>
- Andrews, L. (2014). *Ministering to education: A reformer's report*. Prathian Books.
- Andrews, M. (2013). *The limits of institutional reform in development: Changing rules for realistic solutions*. Cambridge University Press.
- Andrews, M., Pritchett, L., & Woolcock, M. (2017). *Building state capability: Evidence, analysis, action*. Oxford University Press.
- Andersen, S. C., & D. P. Moynihan. (2016). Bureaucratic investments in expertise: Evidence from a randomized controlled field trial. *Journal of Politics*, 78(4), 74), 1032–4104.
- Apreku, S. (2013). *Do and can development evaluations make a difference? The case of the Strategy & Policy Unit Project, Sierra Leone*. Real World African Evaluations in Africa, African Development Bank Group. <https://www.afdb.org/fileadmin/uploads/opev/Documents/Evaluation%20Matters%20March%202013%20Real%20World%20African%20Evaluations%20in%20Africa%20that%20made%20a%20difference%20-%20EN.pdf>
- Bahamas Information Services (BIS). (2018). *Prime Minister's Delivery Unit will be data driven*. Press Release.
- Bandiera, O, Barankay, I. & Rasul, I. (2007). Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *The Quarterly Journal of Economics*, 122(2), 729–773. <https://doi.org/10.1162/qjec.122.2.729>
- Barber, M. (2013). The origins and practice of delivery. *Voices on Society: The Art and Science of Delivery*, 5. McKinsey & Company.
- Barber, M. (2015). *How to run a government: So that citizens benefit and taxpayers don't go*

crazy. Penguins.

Barber, M., Kihn, P., & Motfit, A. (2011, February). Deliverology: From idea to implementation. *McKinsey Our Insights*. <https://www.mckinsey.com/industries/public-sector/our-insights/deliverology-from-idea-to-implementation>

Barber, M., Motfit, A., & Kihn, P. (2010). *Deliverology 101: A field guide for educational leaders*. Sage Ltd.

Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming the English health care system. *Public Administration*, 84(3), 517–538. https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1467-9299.2006.00600.x?saml_referrer

Booth, D. (2012). *Development as a collective action problem: Addressing real challenges of African governance*. Africa Power and Politics Programme.

Brinkerhoff, D. W. (2004). Accountability and health systems: Toward conceptual clarity and policy relevance. *Health Policy and Planning*, 19(6), 371-379.

Buerkli, D. (2015). *Welcome to MindLab-designer of delivery and innovation in Denmark*. Centre for Public Impact. <https://www.centreforpublicimpact.org/welcome-to-mindlab>

Carpenter, D. (2001). *The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies, 1862–1928*. Princeton University Press.

Centre for Public Impact (CPI). (2016). *Cape Town's crusaders: Planning to deliver*. <https://www.centreforpublicimpact.org/insights/cape-towns-delivery-crusaders>

Chambers, V., Wild, L., King, M., & Harris, D. (2012). *Common constraints and incentive problems in service delivery*. ODI.

Chun, Y. H., & Rainey, H. G. (2005). Goal ambiguity and organizational performance in U.S. federal agencies. *Journal of Public Administration Research and Theory*, 15, 529–557.

Delivery Associates. (2018). *Success delivered: How delivery units make a difference for governments and the citizens they serve*. <https://www.deliveryassociates.com/publications/success-delivered>

Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4), 1241–1278.

EuroAid. (2016). *Jordan-Amman: Technical assistance for monitoring and evaluation to the office of the prime minister's delivery unit (PMDU)*. <https://ted.europa.eu/udl?uri=TED:NOTICE:034236-2016:TEXT:EN:HTML>

European Commission. (2018). *Commission of implementing decision*. https://ec.europa.eu/neighbourhood-enlargement/sites/near/files/c_2018_8182_2018_jordan_mod_aap_2015_merged_document.pdf

Executive Mansion. (n.d.). *The history of the president's delivery unit*. <https://www.emansion.gov.lr/2content.php?sub=134&related=45&third=134&pg=sp>

FCDO. (2019). *Education beyond aid (EBA) business case: Summary sheet*. <https://devtracker.fcdo.gov.uk/projects/GB-GOV-1-300517/documents>

Finer, H. (1941). Administrative responsibility in democratic government. *Public Administration Review*, 1(4), 335–350.

Freeguard, G., & Gold, J. (2015). *Data-driven delivery: Lessons from the O'Malley administration of Maryland*. Institute for Government. <https://www.instituteforgovernment.org.uk/publications/data-driven-delivery-lessons-o%E2%80%99malley-administration-maryland>

Friedrich, C. (1940). Public policy and the nature of administrative responsibility. *Public Policy* 1, 3–24. Reprinted in Francis Rourke, Ed., *Bureaucratic power in national politics*, 3rd ed., 1978. Little, Brown.

Gerrish, E. (2016). The impact of performance management on performance in public organizations: A meta-analysis. *Public Administration Review*, 76(1), 48-66.

Global Partnership for Education (GPE). (2020). *Transforming education service delivery in Sierra Leone: From evidence to action*. <https://www.globalpartnership.org/blog/transforming-education-service-delivery-sierra-leone-evidence-action>

Gold, J. (2014). *International Delivery: Centres of government and the drive for better policy im-*

plementation. <https://www.instituteforgovernment.org.uk/publications/international-delivery-centres-government-and-drive-better-policy-implementation>

Gold, J. (2017). *Tracking delivery: Global trends and warning signs in delivery units*. Institute for Government. <https://www.instituteforgovernment.org.uk/sites/default/files/publications/Global%20Delivery%20report.pdf>

Goldfinch, S., Derouen Jr., K., & Pospieszna, P. (2013). Flying blind? Evidence for good governance public management reform agendas, implementation and outcomes in low income countries. *Public Administration and Development*, 33, 50-61.

Gruening, G. (2001). Origin and theoretical basis of New Public Management. *International Public Management Journal*, (4)1, 1-25.

Hammerschmid, G., Van de Walle, S., Andrews, R., & Bezes, P. (2016). *Public administration reforms in Europe: The view from the top*. Edward Elgar.

Hart, C. (2017). *The Albania Delivery Unit: A case study on accountability in action*. https://albania.growthlab.cid.harvard.edu/files/albaniagrowthlab/files/hart_2017_accountability_in_action_-_albanian_delivery_unit.pdf?m=1527708899

Harrison, T. (2016, July). *The role of the centre in driving government priorities: The experience of 'delivery units'* (Working paper). Oxford Policy Management. <https://www.opml.co.uk/files/Publications/corporate-publications/working-papers/wp-role-centre-driving-government-priorities.pdf?noredirect=1>

Hood, C., & Dixon, R. (2010). The political payoff from performance target systems: No-brainer or no-gainer? *Journal of Public Administration Research and Theory*, (20)2, i281-i298.

Implementation Support and Follow-Up Unit (ISFU). (2017). *Annual report*. https://isfu.gov.om/Annual_Report_2017_EN.pdf

ISFU. (2017). *Bylaw number 33 for the establishment of the ISFU*. <https://qanoon.com/p/2018/rd2018033/>

Iyer, D. (2012). *Improving coordination and prioritization: Streamlining Rwanda's National*

Leadership Retreat, 2008-2011. Innovations for Successful Societies, Princeton University.
https://successfulsocieties.princeton.edu/sites/successfulsocieties/files/Rwanda%20COG_Final%20ToU_1.pdf

Kaplan, Robert S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 70(1), 71–9.

Kaplan, Robert S., & Norton, D. P. (1993). Putting the scorecard to work. *Harvard Business Review*, 71(5), 134–47.

Khan, A. Q., Khwaja, A. I., & Olken, B. A. (2016). Tax farming redux: Experimental evidence on performance pay for tax collectors. *Quarterly Journal of Economics*, 131(1), 219-271. <http://eprints.lse.ac.uk/66265/>

Kohli, J., Moody, C., & Buskey, M. (2016). *How should a delivery unit be designed?* Delivering Results Series. Deloitte. <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/public-sector/us-fed-delivering-results-series.pdf>

Kuwait Voluntary National Review (KVNR). (2019). *Report on the Implementation of the 2030: Agenda to the UN high-level political forum on sustainable development*.
https://sustainabledevelopment.un.org/content/documents/23384Kuwait_VNR_FINAL.PDF

Laboratorio Para La Ciudad (LABCD). (n.d.) *Experimentacion y creatividad en la Ciudad de Mexico (Archivo 2013-2018)*. <https://labcd.mx/>.

Lafuente, M. & Gonzalez, S. (2018). *Do delivery units deliver?: Assessing government innovations* (Technical Note No. IDB-TN-1431). Inter-American Development Bank, Innovations in Citizen Services Division.

London Borough of Barking and Dagenham. (2017). *Prevent strategy and delivery plan*.
<https://www.lbbd.gov.uk/sites/default/files/attachments/201801-prevent-strategy-and-delivery-plan.pdf>

McKay, S. (2017). *Quality improvement: Deliverology*. Carnegie Commons Bog. Carnegie Foundation for the Advancement of Teaching. <https://www.carnegiefoundation.org/blog/quality-improvement-deliverology/>

Ministry of Education (MoE), Ecuador. (n.d.) *Planta central* (Spanish). <https://educacion.gob.ec/wp-content/uploads/downloads/2014/03/3.-COORDINACION-GENERAL-DE-SECRETARIA.pdf>

Ministry of Education (MoE), Jordan. (2018). *Education strategic plan: 2018-2022*. http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Amman/pdf/ESP_English.pdf

Ministry of Education (MoE), Liberia. (2018). *Ministry of Education: Education Delivery Unit vacancy announcement*. <https://emansion.gov.lr/doc/ECE%20S.pdf>

Ministry of Information, Culture, and Tourism. (2012). *Lifting Liberians...150 day action plan implementation report*. <http://www.liberianembassyus.org/uploads/PDFFiles/150-Day%20Report.pdf>

Mourshed, M., Chijoke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. McKinsey & Company. https://www.mckinsey.com/~media/McKinsey/Industries/Public%20and%20Social%20Sector/Our%20Insights/How%20the%20worlds%20most%20improved%20school%20systems%20keep%20getting%20better/How_the_worlds_most_improved_school_systems_keep_getting_better.pdf

Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Georgetown University Press.

Moynihan, D., Fernandez, S., Soonhee Kim, K. M., LeRoux, S. J. Piotrowski, B. E., & Wright, K. Y. (2008). Performance regimes amidst governance complexity. *Journal of Public Administration Research and Theory*, 21, i141-i15.5

National Reconstruction Authority. (2015). *Central Level Project Implementation Unit (Education)*. <http://www.moepiu.gov.np/en/>

NCHRD. (2016). *National Human resource development strategy 2016-2025*. National Committee for Human Resource Development. <https://www.hrd.jo/nationalstrategy>

O'Malley, M. (2019). *Smarter government: How to govern for results in the information age*. Esri Press.

PEMANDU. (n.d.). *Turning national visions and plans into big fast results.*

<https://pemandu.org/government/>

PMDU The Bahamas. (2019). *PMDU annual report.* https://mofstii.gov.bs/wp-content/uploads/2020/05/PMDU-Annual-Report-2018_2019.pdf

Pollitt, C., & Bouckaert, G. (2004). *Public management reform: A comparative analysis.* Oxford University Press.

Presidential Delivery Unit. (n.d.). *Big four.* <https://big4.delivery.go.ke/>

Pritchett, L. (2015). *Creating education systems coherent for learning outcomes* (Working Paper 15/005). RISE Programme. www.riseprogramme.org/publications/rise-working-paper-15005-creating-education-systems-coherent-learning-outcomes

Qarout, D. (n.d.). *Delivering education in Jordan: Multiple accountabilities* (Working paper). DeliverEd.

Radin, B. A. (2006), *Challenging the performance movement.* Georgetown University Press, Georgetown University.

Rainey, H. G., & P.SteinbRaineyauer. (1999). Galloping elephants: Developing elements of a theory of effective government organizations. *Journal of Public Administration Research and Theory*, 9, 1–32.

Rasul, I., & D. Rogger. (2018). Management of bureaucrats and public service delivery: Evidence from the Nigerian Civil Service. *Economic Journal*, 128, 413–446.

Rasul, R., Rogger, D., & Williams, M. J. (2020). Management, organizational performance, and task clarity: Evidence from Ghana's Civil Service. *Journal of Public Administration Research and Theory*.

Rose-Ackerman, S. (1986). Reforming public bureaucracy through economic incentives? *Journal of Law, Economics and Organization*, 2, 131–161.

Scharff, M. (2012). *Delivering on a presidential agenda: Sierra Leone's Strategy and Policy Unit, 2010-2011.* Princeton University, Innovation for Successful Societies. https://successfulsocieties.princeton.edu/sites/successfulsocieties/files/Sierra%20Leone%20President%20ToU_1_

0.pdf

Scharff, M. (2013). *Translating vision into action: Indonesia's Delivery Unit, 2009-2012*. Innovations for Successful Societies. Princeton University. https://successfulsocieties.princeton.edu/sites/successfulsocieties/files/Policy_Note_ID224.pdf

Schick, A. (1998). Why most developing countries should not try New Zealand reforms. *The World Bank Research Observer*, 13, 123–131.

Shostak, R., Watkins, J., Bellver, A., & John-Graham, I. (2014). *When might the introduction of a delivery unit be the right intervention?* World Bank Group-Governance and Public Sector Management.

Simson, R. (2013). *Unblocking results - Case study: The Africa Governance Initiative in Sierra Leone*. Overseas Development Institute. www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8412.pdf

Simon, W. (1983). Legality, bureaucracy, and class in the welfare system. *Yale Law Journal*, 92, 1198–1269.

Tierney, M. J., Nielson, D. L., Hawkins, D. G, Roberts, J. T., Findley, M. G., Powers, R. M., Parks, B., Wilson, S. E., & Hicks, R. L. (2011). More dollars than sense: Refining our knowledge of development finance using AidData. *World Development*, 39(11), 1891-1906.

Todd, R. (2014). *Delivery units: Can they catalyze sustained improvements in education service delivery?* Cambridge Education.

Todd, R. (2018). *UNICEF Think Piece series: Accountability*. UNICEF Eastern and Southern Africa Regional Office.

Todd, R., & Waistell, D. (forthcoming). *Overview of the delivery approach*. Cambridge Education report for the Bill and Melinda Gates Foundation.

United Nations Development Programme (UNDP). (2019). *International consultant to support the establishment of a 'delivery unit'*. https://jobs.undp.org/cj_view_job.cfm?cur_job_id=85219

Williams, M. J., Leaver C., Mundy K., Mansoor, Z., Qarout, D., Asim, M., Bell, S., & Bilous, A. (2020, June). *Delivery approaches to improving policy implementation: A conceptual framework*. DeliverEd Working Paper. <https://bit.ly/DeliverEdFramework>

Wilson, J. (1989). *Bureaucracy: What government agencies do and why they do it*. Basic Books.

World Bank. (2003). *Jordan - Education reform for Knowledge Economy Program Project* (English).

<http://documents.worldbank.org/curated/en/239641468774926089/Jordan-Education-Reform-for-Knowledge-Economy-Program-Project>.

World Bank. (2004). World development report 2005: Making services work for poor people. *Development and Change*, 36(6), 1263-266. <https://openknowledge.worldbank.org/handle/10986/5986>.

World Bank. (2014). *Morocco - Linking education and accountability for development project: Small RETF grant* (English).

<https://documents.worldbank.org/en/publication/documents-reports/documentdetail/734051467990987111/morocco-linking-education-and-accountability-for-development-project-small-retf-grant>

World Bank Group. (2017). *Driving performance from the center: Malaysia's experience with PEMANDU*. <http://documents1.worldbank.org/curated/en/318041492513503891/pdf/114270-WP-PUBLIC-13-4-2017-13-0-58-WorldBankReportPemanduFAFULLWeb.pdf>

World Bank. (2020). *New World Bank country classifications by income level: 2020-2021*. <https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2020-2021>

Appendices

Appendix A: List of countries covered in the mapping

The list of countries was derived from the World Bank's database in 2020, excluding disputed or occupied territories.

sr. no	county name	country code
1	Afghanistan	AFG
2	Albania	ALB
3	Algeria	DZA
4	Andorra	AND
5	Angola	AGO
6	Antigua and Barbuda	ATG
7	Argentina	ARG
8	Armenia	ARM
9	Australia	AUS
10	Austria	AUT
11	Azerbaijan	AZE
12	Bahrain	BHR
13	Bangladesh	BGD
14	Barbados	BRB
15	Belarus	BLR
16	Belgium	BEL
17	Belize	BLZ
18	Benin	BEN
19	Bhutan	BTN
20	Bolivia	BOL
21	Bosnia and Herzegovina	BIH
22	Botswana	BWA
23	Brazil	BRA
24	Brunei	BRN
25	Bulgaria	BGR
26	Burkina Faso	BFA
27	Burundi	BDI
28	Cameroon	CMR
29	Canada	CAN
30	Cape Verde	CPV
31	Cambodia	KHM
32	Central African Republic	CAF
33	Chad	TCD
34	Chile	CHL
35	China	CHN
36	Colombia	COL
37	Comoros	COM

38	Congo, Democratic Republic of the	COD
39	Congo, Republic of the	COG
40	Costa Rica	CRI
41	Cote d'Ivoire	CIV
42	Croatia	HRV
43	Cuba	CUB
44	Cyprus	CYP
45	Czech Republic	CZE
46	Denmark	DNK
47	Djibouti	DJI
48	Dominica	DMA
49	Dominican Republic	DOM
50	East Timor (Timor-Leste)	TLS
51	Ecuador	ECU
52	Egypt	EGY
53	El Salvador	SLV
54	Equatorial Guinea	GNQ
55	Eritrea	ERI
56	Estonia	EST
57	Ethiopia	ETH
58	Fiji	FJI
59	Finland	FIN
60	France	FRA
61	Gabon	GAB
62	Georgia	GEO
63	Germany	DEU
64	Ghana	GHA
65	Greece	GRC
66	Greenland	GRL
67	Grenada	GRD
68	Guatemala	GTM
69	Guinea	GIN
70	Guinea-Bissau	GNB
71	Guyana	GUY
72	Haiti	HTI
73	Honduras	HND
74	Hong Kong SAR, China	HKG
75	Hungary	HUN
76	Iceland	ISL

77	India	IND
78	Indonesia	IDN
79	Iran	IRN
80	Iraq	IRQ
81	Ireland	IRL
82	Israel	ISR
83	Italy	ITA
84	Jamaica	JAM
85	Japan	JPN
86	Jordan	JOR
87	Kazakhstan	KAZ
88	Kenya	KEN
89	Kiribati	KIR
90	Korea, North	PRK
91	Kosovo	XKX
92	Kuwait	KWT
93	Kyrgyzstan	KGZ
94	Laos	LAO
95	Latvia	LVA
96	Lebanon	LBN
97	Lesotho	LSO
98	Liberia	LBR
99	Libya	LBY
100	Liechtenstein	LIE
101	Lithuania	LTU
102	Luxembourg	LUX
103	Macao SAR, China	MAC
104	Macedonia	MKD
105	Madagascar	MDG
106	Malawi	MWI
107	Malaysia	MYS
108	Maldives	MDV
109	Mali	MLI
110	Malta	MLT
111	Marshall Islands	MHL
112	Mauritania	MRT
113	Mauritius	MUS
114	Mexico	MEX
115	Micronesia, Federated States of	FSM

116	Moldova	MDA
117	Monaco	MCO
118	Mongolia	MNG
119	Montenegro	MNE
120	Morocco	MAR
121	Mozambique	MOZ
122	Myanmar	MMR
123	Namibia	NAM
124	Nauru	NRU
125	Nepal	NPL
126	Netherlands	NLD
127	New Zealand	NZL
128	Nicaragua	NIC
129	Niger	NER
130	Nigeria	NGA
131	Norway	NOR
132	Oman	OMN
133	Pakistan	PAK
134	Palau	PLW
135	Panama	PAN
136	Papua New Guinea	PNG
137	Paraguay	PRY
138	Peru	PER
139	Philippines	PHL
140	Poland	POL
141	Portugal	PRT
142	Qatar	QAT
143	Romania	ROU
144	Russia	RUS
145	Rwanda	RWA
146	Saint Kitts and Nevis	KNA
147	Saint Lucia	LCA
148	Saint Vincent and the Grenadines	VCT
149	Samoa	WSM
150	San Marino	SMR
151	Sao Tome and Principe	STP
152	Saudi Arabia	SAU
153	Senegal	SEN
154	Serbia	SRB

155	Seychelles	SYC
156	Sierra Leone	SLE
157	Singapore	SGP
158	Slovakia	SVK
159	Slovenia	SVN
160	Solomon Islands	SLB
161	Somalia	SOM
162	South Africa	ZAF
163	South Korea	KOR
164	South Sudan	SSD
165	Spain	ESP
166	Sri Lanka	LKA
167	Sudan	SDN
168	Suriname	SUR
169	Swaziland	SWZ
170	Sweden	SWE
171	Switzerland	CHE
172	Syria	SYR
173	Taiwan	TWN
174	Tajikistan	TJK
175	Tanzania	TZA
176	Thailand	THA
177	The Bahamas	BHS
178	The Gambia	GMB
179	Togo	TGO
180	Tonga	TON
181	Trinidad and Tobago	TTO
182	Tunisia	TUN
183	Turkey	TUR
184	Turkmenistan	TKM
185	Tuvalu	TUV
186	Uganda	UGA
187	Ukraine	UKR
188	United Arab Emirates	ARE
189	United Kingdom	GBR
190	United States	USA
191	Uruguay	URY
192	Uzbekistan	UZB
193	Vanuatu	VUT

194	Venezuela	VEN
195	Vietnam	VNM
196	Yemen	YEM
197	Yugoslavia	YUG
198	Zambia	ZMB
199	Zimbabwe	ZWE

Appendix B: Search Protocol

To identify cases of delivery approaches (DAs), we relied on the following sources (in this sequence):

- a) Existing lists of delivery units, from DeliverEd preliminary review or other published documents
- b) Structured desk-based search and document review, which we conducted for each country and some subnational jurisdictions
- c) References to a DA from documents on another DA, for example when a DA identified in the preceding steps has a website or a donor project report that mentions an earlier DA

Step (b), the structured desk-based search, which was the main component of the search, was conducted via systematic desk-based search of the following sources:

- AidData’s global database of donor projects since 1990 (Tierney et al., 2011), Research Release 3.1. We examined the subset of projects with DAC-CRS codes 15110 and 15111, which are the two codes that are most relevant to public management reforms.
- The World Bank’s database of project documents for 1990-present (using the set of key search terms defined below) (see <https://www.worldbank.org/en/region/lac/projects/all>)
- Google and Google Scholar (using a set of key search terms)
- Oxford University’s SOLO index of e-journals

For gray literature and case studies on delivery units/approaches, we drew on the following databases:

- The Global Delivery Initiative Library database (<http://www.globaldeliveryinitiative.org/search/library/Prime%20Minister%27s%20Delivery%20Uni>)
- Centre for Public Policy Impact, the Public Impact Observatory (<https://www.centreforpublicimpact.org/insights>)

- Inter-American Development Bank (IADB) (<https://www.iadb.org/en/projects>)
- African Development Bank Group (ADB) (<https://www.afdb.org/en?id=4338>)
- ISE Reform Sequencing Tracker (<https://reformtracker.effectivestates.org/>)

We defined a short list of key search terms we used to search each database, which we then searched for each country (and some subnational units) in the world. For example:

- “COUNTRYNAME” + “delivery unit”
- “COUNTRYNAME” + “delivery lab”
- “COUNTRYNAME” + “reform secretariat”
- “COUNTRYNAME” + “monitoring and evaluation unit”
- “COUNTRYNAME” + “implementation unit”

For completeness, we also conducted searches in these same databases for documents related to cases we already knew existed from our prior outreach or knowledge. For example:

- “GHANA” + “Presidential Delivery Unit”

Where a unit may be referred to by multiple names, we were flexible in searching to ensure that we identified the relevant documents, for example through variations in how terms are combined such as:

- “GHANA” + “Ministry of Education” + “Reform Unit”
- “GHANA” + “education” + “Reform Unit”

We then conducted a broader search using general terms to identify cases and sources that were not captured above. Using mainly Google Scholar (but using also with other search tools and specific journals), we searched the following terms in combination with the country name (i.e., “Ghana” + “TERM FROM LIST”):

- *General terms.* Delivery units, delivery labs, president’s implementation unit, implementation unit, Center of Government delivery unit, planning secretariat, National Steering Committee, reform secretariat, Premier’s Implementation Unit, Governor’s delivery unit (or other variation of the unit)

- *Institution names.* Prime Minister’s Office (PMO), Reform Secretariat, Cabinet, Results and Effectiveness Unit, National Project Implementation Unit, Performance Management Unit, Premier’s Office, etc.

We used the above as a provisional list to be used as a starting point, so that coders could also try different terms, terms in combinations, using Boolean operators to eliminate irrelevant results (e.g., using minus signs to eliminate terms that are yielding irrelevant results), and so on. Throughout the search process, we held regular team check-ins in which we discussed and harmonized search strategies to ensure consistency and maximum coverage.

To search for subnational or regional DAs in our six countries in which we conducted systematic subnational mapping (Australia, Canada, China, India, Pakistan, and the United States), the following steps were taken:

- For each country, we identified the political system and the level of decentralization, as well as a list of the biggest states, cities, or provinces.
- We then developed a list of titles used for subnational or regional executives, as well as the organizational names.
- When searching for units at these subnational or regional levels, we searched on “executive title or organization name” + “general term” and on “name of State or Province/City/Region” + “general term”.

Finally, in step (c) cases and sources were compiled from incidental references we saw in documents/websites we reviewed as part of steps (a) and (b).

Appendix C: Coding Protocol and Variables

The coding of the cases relied primarily on publicly available data from secondary sources. The variables that were coded fall under two categories. The first is basic characteristics, including the start and end dates of the DA, the sector(s) in which the approach operated, and the level of government at which the approach was established. The second is design features of the approach, including the time horizon and nature of the goals (outputs vs. outcomes) the approach is trying to achieve, the staffing structure and institutional set-up of the approach, whether the approach received any external funding or technical assistance, and the managerial mechanisms through which the approach sought to improve implementation and bureaucratic performance. Judgements were made by coders based on the available data and triangulated the information across multiple sources to ensure accuracy to the extent possible. The coders aimed to ensure consistency in the coding by meeting on a weekly basis to discuss coding decisions and interpretation of data.

This appendix describes each item that was coded for each case identified by the search protocol. Items are listed in bold, with potential coding options in italics under the bullet points. For many response codes, we also provide an example to help illustrate for readers the type(s) of qualitative information that each coding option corresponds to.

Delivery approach type: Is this a unit-type delivery approach or a structured process?

- *Unit* The delivery approach is structured as a unit or a department within a bureaucracy.
- *Structured process.* The DA is structured as a process that intersects with multiple units or departments within a bureaucracy to leverage key management functions fragmented across more than one bureaucratic structure but does not exclusively rely on a central team or structure.

e.g., Kuwait’s Supreme Council for Planning and Development runs a committee on monitoring and evaluation implementation performance against their Sustainability Development Goals (SDGs) delivery plan. The Committee is comprised of representatives from various line ministries and organizations involved in the delivery process and runs a series of consultation workshops, target review session, and problem-solving discussions around implementation progress and bottlenecks (KNR, 2019). The approach cuts across multiple public agencies and is not housed in a specific bureaucratic structure, but involves a series of activities and routines carried out by different de-

partments, including monitoring, problem solving, and routine reporting to political officials (KNR, 2019).

Sector: In which sector/s does the delivery approach operate? This entails the sectors linked to the priority goals the delivery approach is tasked with overseeing.

- *Education sector.* The DA is embedded exclusively within the education sector, monitoring priority goals that link to educational outputs and/or outcomes and engaging with stakeholders involved in education service delivery.

e.g., Liberia’s Ministry of Education adopted its own Education Delivery Unit (EDU), which is tasked with ensuring the implementation of the sector’s 3-year plans (MoE, *Liberia*, 2018).

- *Other single sector.* The DA is embedded exclusively within a single sector, monitoring priority goals that link to that sector’s outputs and/or outcomes and engaging with stakeholders involved in service delivery under that sector’s remit.

e.g., The UK’s Department for Defense established its own Ministerial Delivery Unit, focusing on sector-specific national priority goals set by the Secretary of State for Defense (Gold, 2017).

- *Multiple sectors.* The DA operates across multiple sectors, monitoring priority goals that link to more than one sector’s outputs and/or outcomes and engaging with stakeholders involved in service delivery across several sectors, line ministries and organizations.

e.g., Jordan’s Prime Minister’s Delivery Unit (PMDU) operates across all sectors within government, coordinating with over 120 line ministries to ensure the implementation of the national priority reforms (EuroAid, 2016; NCHRD, 2016).

Goal time horizon: What was the time horizon for the explicit and/or implicit goals of the delivery approach? This refers to the time period over which the goals are intended to be implemented.

- *Mostly short-term.* Policies and goals that are targeted should be achieved in the next 1-2 years

e.g., Liberia’s President’s Delivery Unit oversaw the implementation of President Sirleaf’s 150 day plan, which laid out short-term goals such as constructing 1000 sanitation facilities (Ministry of Information, Culture, and Tourism, Liberia, 2012).

- *Mostly medium-term.* Policies and goals that are targeted should be achieved in the next 3-5 years.
 - e.g., Sierra Leone’s Strategy and Policy Unit was tasked with implementing the President’s Agenda for Change (2007-2012), a vision document that outlined goals for improvements in energy, transportation, health, and education (Scharff, 2012).
- *Mostly long-term.* Policies and goals that are targeted should be achieved in the next 6 or more years
 - e.g., Kenya’s Presidential Delivery Unit is tasked with ensuring the implementation of national priorities over 5-10 years, including incrementally raising the contribution of manufacturing and agriculture to the country’s GDP (PDU, n.d.).
- *Mixed.* A combination of short, medium, and/or long-term goal time horizons.
 - e.g., Jordan’s Results and Effectiveness Unit monitored priorities with time horizons that spanned from 1 year (e.g., establishment of a National Center for Curriculum Development) to 6 years and more (e.g., universalizing enrollment in KG2.) (NCHRD, 2016).

Goal type: What was the nature of the goals being monitoring by the delivery approach? Were they mostly focused on achieving outcomes at the level of the citizen/client or were they focused on ensuring outputs from various policies, programs/interventions, and reform are produced across each level of administration (including frontline) and delivery chain?

- *Mostly outcomes-based goals.* The goals adopted by the DA are mostly focused on achieving outcomes at the frontline and citizen/client level. Outcomes are defined as the impacts or benefits of policy delivery for service beneficiaries (e.g., improved test scores, increased job opportunities, higher income for workers, etc.)
 - e.g., Kenya’s Presidential Delivery Unit is focused on ensuring effective implementation in four priority areas: manufacturing, affordable housing, universal health coverage, and food security. The majority of the goals set for each domain are outcome focused. Examples of these outcomes include the reduction of food malnutrition among children under age 5 by 27% and increasing daily income for farmers by 34% (PDU, n.d.)
- *Mostly outputs-based goals.* The goals adopted by the DA are mostly focused on achieving outputs across the delivery chain, including central government, subnational, and frontline administrations. Outputs are defined as actions taken by government agen-

cies or employees, such as the programs, trainings, materials, products and services produced from a policy, reform, or intervention.

e.g., HRD Strategy priorities monitored by the Results and Effectiveness Unit in Jordan in the short to medium term include mainly outputs. The outputs include: establishment of a Curriculum Center, establishment of an Initial Teacher Education program at the postgraduate level and increase capacity for higher student enrolment, increasing enrolment of children in KG2, increasing public school capacity for KG2 enrolment, improving governance and cross-sectoral coordination of the Technical and Vocation Education and Training (TVET) sector.

- *About equal* The goals adopted by the DU/approach are a more or less equal combination of both outcomes and outputs.

e.g., The Prime Minister’s Delivery Unit (PMDU) in the Bahamas identified 6 priority areas: Over-the-Hill Rejuvenation Project, Land Reform, Energy Reform and Infrastructure, Ease of Doing Business, Education, and Safety and Security (BIS, 2018). Within each of these domains, the goals are more or less equally focused on both outcomes and outputs. For outcomes, some examples include improving student literacy and promoting evidence-based pedagogy in the classroom. For outputs, some examples include ensuring all public schools have internet connectivity and universalising preschool education (BIS, 2018).

Staffing: Were staff recruited for the delivery approach civil servants, consultants from the private/NGO sectors or a combination of both?

- *Civil service staff.* Staff entirely recruited from the civil service.

e.g., UK’s Department for Education Delivery Unit staff are entirely recruited from within the civil service

- *Non-civil service staff.* Staff consists of consultants, technical assistants, or other external hires working outside the mainstream civil service.

e.g., An evaluation of Sierra Leone’s Strategy and Policy Unit’s functions between 2008 and 2010 commissioned by the UNDP and conducted by the African Center for Economic Transformation (ACET) found that the Unit lacked civil service representation. The staff of the SPU were recruited through a competitive process that involved public advertisement, assessments, and screening interviews to identify experts. This staffing arrangement reportedly caused tensions between ministry staff and the SPU members, which ultimately led to inefficient ways of working and slow progress towards

intended targets. An evaluation published by the African Development Bank Group in 2013 suggests that civil servants were offered the opportunity to be seconded to the SPU to mitigate the tension between SPU staff and civil servants (ACET, 2010).

- *Mixed*. Staff include both civil servants and hired consultants from outside the mainstream civil service.

e.g., The Education Beyond Aid (EBA) Programme Business Case published by the FCDO in 2019 provided detailed information about the establishment of the Reform Secretariat in Ghana, as well as the kind of technical assistance they will be receiving. The Reform Secretariat is staffed by civil service personnel and five advisers recruited from outside the civil service. The advisers provide technical support in the team’s collection and analysis of data, monitoring of implementation progress, and brainstorming solutions for implementation challenges (FCDO, 2019).

Structure: What was the institutional set-up of the delivery approach? Was it hosted in preexisting bureaucratic structures/processes within government or new/reorganized structures?

- *Using preexisting bureaucratic structures and processes*. Use of preexisting structures within the bureaucracy to execute the key functions of the DA.

e.g., When Jordan adopted a new DA in the Education sector in 2018, the Ministry of Education used the preexisting Development Coordination Unit (DCU) to operationalize the approach and tasked the DCU with a more active role in monitoring all sector priorities, not just donor-funded projects or disbursement-linked indicators outlined by donors as it was originally designed for. The DCU became responsible for ensuring the implementation of the Education Sector Plan (2018-2023) in 2018 and took on a broader mandate to provide oversight of delivery across the MoE, coordinating with various departments, reporting to the Minister on progress and implementation challenges, and providing support to the Technical Working Groups as needed (World Bank, 2003; MoE, *Jordan*, 2018).

- *Using new or reorganized bureaucratic structures and processes*. An entirely new bureaucratic structure or a hybrid of existing and new bureaucratic structures, or a reorganizing or shuffling of preexisting bureaucratic structures into a novel configuration.

e.g., The Indonesian President Yudhoyono established an entirely new unit within the presidency to track the President’s priority reforms. He also exercised discretion

in hiring and compensating staff to ensure that the new unit attracted high calibre staff and introduced innovation to the bureaucracy which disrupted the status quo. The delivery unit emerged as a new structure within the bureaucracy that was not originally part of the organizational chart (Scharff, 2013).

- *Using structures outside the government bureaucracy.* Creation of a new unit that exists outside or in parallel to existing structures of bureaucratic hierarchy and resource allocation.

Level: The level of government at which the DA is set up. The DA may operate at multiple levels, but this variable was coded based on the institutional host of the DA and the headquarters of its operations.

- *National-Center of Government.* The DA sits primarily at the top-most level of political and bureaucratic authority, such as a president or prime minister’s office.
- *National-Ministerial.* The DA sits primarily at the national-level sector ministry or other national-level agencies, such as an education ministry.
- *Subnational.* The DA sits primarily at the government entity that exists below the national level. This could be states, provinces, regions, districts, or municipalities.
- *Frontline.* The DA sits primarily at the level where service delivery takes place, i.e., where the front-line providers of the service directly engage with citizens. This would, for example, comprise schools in the education sector or basic health units in the health sector.

External funding: Any funding from external agencies not working within the host country government to support the set-up and operationalization of the DA.

- *Yes.* The DA has received significant external funding from development banks, NGOs, and/or the private sector to support its establishment and/or its operations. The funds could be allocated towards staff, training, and other expenses needed for the delivery unit/approach to exist and function.

e.g., The FCDO provided funding for technical assistance to staff working at the Reform Secretariat in Ghana, building their capacity in monitoring and evaluation, creating data systems and analyzing data to inform decisions and implementation (FCDO, 2019).

- *No.* Based on the information available, the DA did not appear to receive significant external funding from development banks, NGOs, and/or the private sector.

Technical assistance: Technical support received by the DA for advising on the institutional framework, operations, and terms of reference of the approach, as well as supporting the operationalization of the operation. Technical assistance also includes the training provided to staff working in the DA by external consultant or vendors.

- *Yes.* The DA received technical assistance in the form of training, seconded staff, and/or advice around the design and functioning of the DA.

e.g., In Tbilisi, Georgia, the UNDP published a job posting for consultants to work on supporting the establishment and operations of a delivery unit within the Ministry of Internally Displaced Persons from the Occupied territories, Labour, Health and Social Affairs of Georgia. The scope of work published in the job posting identified 4 main areas of technical support: conducting a situation analysis of the Ministry and sector to identify the needs for the establishment of a delivery unit, development an institutional framework for the delivery unit and determining the functions and job descriptions for the delivery unit staff, developing a workplan and roadmap for the set-up of the DU, and providing training to the delivery unit staff in the areas of monitoring and communications.

- *No.* Based on the information available, the DA did not appear to receive any technical assistance in the form of training, seconded staff, and/or advice around the design and functioning of the DA.

Managerial mechanisms: Managerial functions and routines to improve policy implementation and bureaucratic performance.

- *More accountability- and incentive-driven mechanisms.* The DA relies more heavily on top-down incentives and accountability as its key function.

e.g., In 2008, Sierra Leone’s President Koroma set up the Strategy and Policy Unit (SPU) with the aim of instituting a stronger structure and greater accountability (ISS, 2011). The SPU’s managerial approach focused on more accountability-and-incentive driven routines in 2008, although after an external evaluation of the SPU in 2010 the unit was advised to include problem-solving elements as well (Simson, 2013). The accountability routines included: performance contracts with the Minister, tracker sheets (or roadmaps) for specific priorities that outlined who was responsible for what

to track progress including RAG ratings, and monthly meetings (called stocktakes) at the president's office to review the status of each flagship project. The top-down accountability was strong. For example, the President signalled the seriousness of accountability by dismissing a Minister after learning during a stocktake meeting that this Minister was not meeting set targets (Simson, 2013).

- *More problem-solving and organizational learning-driven mechanisms.* The DA relies more heavily on horizontal and vertical convening, collaboration, and problem-solving as its key function.

e.g., The Commissioner's Delivery Unit (CDU) in the state of Kentucky, United States, was designed to primarily research the effectiveness of the administration's delivery and adapt its strategy and delivery plans to meet the needs of their beneficiaries. The approach implemented Design-Based Implementation Research (DBIR), where research on the implementation process informed the various iterations of the implementation plan. The approach also relied on Networked Improvement Communities (NIC) to identify and solve complex problems in implementation both systematically and scientifically, using data to drive the improvement process. NICs were also leveraged for productive problem solving, capitalising on expertise of different stakeholders with a common purpose of achieving similar outcomes. This involved information sharing, consultation, and regular discussions for iterative problem solving (McKay, 2017).

- *About equal.* The DA exhibits a more or less equal combination of both accountability and incentive-driven managerial mechanisms and problem-solving and organizational learning-driven mechanisms.

e.g., The PMDU in the Bahamas was designed to drive results-oriented implementation by focusing heavily on data to analyze progress and challenges, thus ensuring civil servants' engagement in continuous problem solving throughout the implementation process. The Unit additionally leveraged high-level meetings and monitoring routines to signal accountability for results. These routines involved: periodic data collection and analysis, routine Delivery Update Meetings (between Delivery Unit and Stakeholder Delivery Teams, monthly written reports summarising updates on implementation for each priority area, Stocktake meetings between the Prime Minister and Accountable Leaders on a quarterly basis, and regular review of delivery plans based on data (PMDU, 2019).

Income level. We rely on the World Bank's definition of country income levels, which

are divided into four categories: high income, upper-middle income, lower-middle income, and low income (World Bank, 2020).

- *High-income*. GNI per capita over USD 12,535
- *Upper-middle income*. GNI per capita between USD 4,046 and USD 12,535
- *Lower-middle income*. GNI per capita between USD 1,036 and USD 4,045.
- *Low-income*. GNI per capita of less than USD 1,036.