



# PART 1

## Overview





# CHAPTER 1

## Introduction to *The Government Analytics Handbook*

*Daniel Rogger and Christian Schuster*

### THE TRANSFORMATIVE POWER OF GOVERNMENT ANALYTICS

In 2016, Kim Wells, a senior member of the US federal government's Federal Employee Viewpoint Survey (FEVS) team—perhaps the world's most famous survey of public servants—had an appointment with an imposing ex-marine who had entered the public service as a manager. The marine still had a soldier's physicality about him as he entered the conference room with Kim, and you could see that he brought his military management style with him to his duties. For so many members of the US military, the idea of excellence is fundamental to how they see themselves and their work. His identity was rooted in the idea that he was an outstanding manager. And if you had asked him, he would have woven a narrative of success against all odds in the work that he and his team were doing.

Yet he was failing: failing to create a work environment for his staff in which they felt engaged, mentored, and safe. Kim had surveyed his entire team, giving each of them a chance to provide feedback and for that feedback to be compared with the experiences of other officials working under different managers. And the truth was that this burly ex-marine was failing his team, himself, and his country. He broke down in tears in front of Kim. His view of himself had been confronted by the survey data that gave his staff a voice they could not have had otherwise. He knew he needed to change and improve how he managed his team—and the survey data told him exactly how to go about doing it.

At roughly the same time, but more than four thousand miles to the south, Brazil's federal government was heading for financial catastrophe. Under the existing pay and pensions regime that compensated federal employees, wage costs were about to skyrocket. Seeing the impending danger through the dense wording of public contracting law was daunting. Mirian, a member of the Ministry of Economy, suspected something was wrong, but couldn't put a finger on what lay ahead. So, Mirian had a team calculate what the future of payroll and pensions looked like for every individual in the federal government under the existing regime. The danger suddenly seemed very real. As wages skyrocketed, funding for other inputs to government services would become unaffordable. Services would have to be stopped. Fortunately, Mirian had also asked the team to model other feasible scenarios. These cases gave the government the means to negotiate with politicians and other stakeholders and pass legislation to change compensation rules in time to avert catastrophe.

Four thousand miles east, the government of Nigeria had received debt relief from the Paris Club group of creditors in 2005 worth US\$18 billion. Many Nigerians wanted to know what would happen to those funds, including Amina Mohammed, a northern Nigerian who was almost invariably clothed in traditional dress and had a background in both engineering and civil society advocacy. The president asked Amina to join his team, and over the next few years she built one of the world's most innovative public sector tracking systems to follow the financial gains of debt relief through government.

From her office in the Presidency, Amina tracked every naira of those funds, combining budget, program, and audit data systems and sending teams to visit every project site. It was truly frontier analytics, showing where and how badly the government was failing. Some organizations fulfilled their commitments completely. Others did not. In 2006, for instance, the Ministry of Water Resources received US\$475 million to develop water infrastructure across the country. In return, it produced nothing. The ministry's officials seemed busy and budgetary releases were made. But when Amina's teams visited the sites that had been allocated funds all across the country, they could not find a single project that had been completed. Amina took this evidence to the president and won the political and bureaucratic space she needed to create new ways of spending government resources, such as a grants scheme to state governments that only paid out if water infrastructure was actually produced.

This book, *The Government Analytics Handbook*, is about enabling individuals like Kim, Mirian, and Amina to change their governments for the better. It draws on a moment in history when the world is capitalizing on innovations in measurement, data collection, and data analysis at an unprecedented scale. Never before has the world been able to build a richer picture of the realities of the public sector. The question for each and every public sector official, manager, and leader is what they are going to do with this revolution. How governments collect, analyze, and use microdata to improve the administration of government—or undertake what this *Handbook* calls government analytics—will determine how effective they are in this new world.

Government analytics can help solve big issues in public administration—governmentwide administrative challenges, as in the case of future fiscal liabilities from Brazil's payroll. But as important it can also help government organizations improve in small ways, addressing specific management challenges in specific teams in specific government organizations, as with the example of the former US marine. When small improvements happen across thousands of teams inside government—as enabled by regular governmentwide employee surveys, for instance—even small changes can transform government.

What do we mean by government analytics? It is the repurposing of administrative and survey data from within government to improve the way government functions. It uses microdata to diagnose the inputs, management practices, processes, outputs, or outcomes in public sector organizations, units inside such organizations, and/or public administration as a whole. These diagnoses can pinpoint how well government is functioning—or not. Microdata provide information about the characteristics of individual people or entities such as individual officials or departmental units in the case of government, or households, business enterprises, or farms in the case of the private sector. Such data can measure and study relationships among phenomena at a very granular scale, such as how the management practices of individual public service managers affect the productivity of their teams. Microdata can come from a range of sources: unit-level data obtained from sample surveys, wider censuses, and general administrative systems.

Government analytics is not restricted to governments with more advanced information technology (IT) platforms and employee records, like Brazil and the United States. Instead, it has been of use to governments around the world and at all stages of administrative development, as illustrated by Amina's efforts in Nigeria. We have had the good fortune of being collaborators and witnesses to many such improvements. Government analytics has, for instance, led to more rigorous merit recruitment procedures in Kosovo's government, better employee onboarding in Chile's government, staff mentoring in Ethiopia's government, higher-quality public service training in Ghana and Nepal, improved quality of management in Croatia's government, and better public procurement practices in Romania and Uruguay. The list goes on. Many more examples are contained across the 30 chapters of this *Handbook*.

Although many instances of government analytics are being carried out at an individual level, there is a lack of systematic practice in governments as a whole. This means that governments are missing out on the potential insights available to them for improving their public administrations at scale. This, in turn, means that money is being left on the table. Public revenues that could be spent more efficiently, with greater impact on the welfare of citizens, are simply not being spent as well as they could. It is time to pick up those funds and use them for a better society.

How can this *Handbook* help? By showcasing how effective and low cost analytics can be, the hope is that more governments will undertake analytics of their own administrations, and in a more systematic way. To make meaningful progress, we need to change the way we approach government, and this shift should reflect a broader change in what we expect to know about state institutions. Analytics have come to dominate discussions of many other spheres of life, and they should play a more significant role in efforts to strengthen the state.

Beyond any single government, there is a lack of systematic evidence on how to do analytics in a rigorous manner, and there are few carefully constructed global comparisons available. As a result, different governments tend to follow diverging practices even when undertaking similar analytics, limiting their ability to use objective benchmarks from other settings. For instance, as shown in chapter 18, different governments ask different questions in their employee surveys to measure the same concepts.

This *Handbook* aims to fill this gap. It presents frontier evidence and practitioner insights on how to leverage data to strengthen public administration. Across 30 chapters, it shows ways to transform the ability of governments to take a data-informed approach to diagnose and improve how public organizations work. An accompanying website contains tools for analytics, which enable readers to immediately apply insights from the *Handbook* in their own work ([www.worldbank.org/governmentanalytics](http://www.worldbank.org/governmentanalytics)). The *Handbook* covers many sources of microdata, ranging from administrative data, such as payroll, procurement, case, text, and human resources management information system (HRMIS) data; to public servant survey data; to data coming from external assessments, such as citizen and household surveys, or anthropological diagnostics of public administration. Methodologically, it covers both traditional qualitative and quantitative methods, as well as newer approaches, such as machine-learning diagnostics of unstructured text records from governments. To our knowledge, this is the first and most comprehensive volume of this kind.

## THE HIGH STAKES OF GOOD GOVERNMENT ANALYTICS

In their magisterial review of the formation of the state, Acemoglu and Robinson (2019, 341) note that “bureaucracy is vital to state capacity.” Whether a country’s laws and policies will indeed be implemented is determined by the quality of its public administration. Extensive research confirms that the quality of government administration affects institutional quality, safety guarantees, education opportunities, health care provision, and ultimately, the welfare of society and the economy (see, among others, Besley et al. 2022; Dahlström and Lapuente 2022; Finan, Olken, and Pande 2015; Pepinsky, Pierskalla, and Sacks 2017; Wilson 1989).

Significant opportunities exist for improving these outcomes through better administration, a variety of studies around the world show. For example, in the Russian Federation, researchers found that the quality of bureaucrats and their organizations accounted for two-thirds of the variation in the cost of public procurement contracts (Best, Hjort, and Szakonyi 2017). By reducing the prices paid by the worst-performing 25 percent of procurement agents to no more than that paid by the other 75 percent of agents, the Russian government could save approximately US\$10 billion each year—a sum equivalent to about 15 percent of Russia’s total public health care spending. Similarly, improving the quality of management in government organizations in Nigeria by a standardized unit could raise by 32 percent the likelihood

that a physical infrastructure project is completed (Rasul and Rogger 2018). In Italy, reassigning managers to place the best managers in the largest offices would boost productivity in social security claim processing by at least 7 percent (Fenizia 2022). In Pakistan, offering the best-performing tax collectors their top choice of job posting would increase tax revenue by 40 percent (Khan, Khwaja, and Olken 2019). Conversely, poor management of the administration leads to worse outcomes. In Brazil, politicized turnover of public personnel, including at schools, significantly lowers student learning (Akhtari, Moreira, and Trucco 2022). Corruption in education funds is associated with a 65 percent increase in dropout rates from schools (Ferraz, Finan, and Moreira 2012).

The essence of these studies is that good management in government can lead to significant and rapid improvements in government performance. Conversely, when public administration is weak or inefficient, programs and policies are far more likely to fail. The size of the findings of recent evaluations of public administration reforms indicates that there is perhaps no more effective means of improving public policy than by strengthening the quality of public administration.

The sheer scale of most countries' public administrations makes their quality important. The cost of wages for public sector employees is approximately 10 percent of gross domestic product (GDP) across the world, not even counting public sector pensions (World Bank 2020; World Bank Group 2019). That is a significant portion of the economy to ensure is managed effectively. Similarly, the assets the public administration manages directly are large. Across the world, public procurement of goods and services accounts for roughly 12 percent of GDP (Bosio and Djankov 2020). As the study of Russian procurement suggests, ensuring that governments are paying appropriate prices for these purchases would yield much more money to spend on health care or other welfare improvements.

A direct consequence of the size of government is its influence over the rest of the economy. Globally, the public sector makes up 38 percent of formal employment (World Bank 2020). As such a large employer, it plays an influential role in the wider labor market, particularly for tertiary educated workers (Somani 2021). The same can be said for the prices it pays for the goods it procures, and the stimulus it induces when it builds infrastructure or regulates business. So even if interest is solely in improving the private sector, government analytics matters.

Yet with so much at stake, and such large margins for improvement to exploit, governments everywhere are not exploiting modern analytics (World Bank 2016; World Bank 2021). Although governments worldwide have invested heavily in digitizing their administrative work—dedicated digital government institutions have been established in 154 economies—few have systematized the use of the resulting records into data to strengthen the way they work, the World Bank's GovTech Maturity Index reveals (World Bank 2022). As noted, analytics is not necessarily dependent on such digitization, but it is illustrative of wider commitments to analytics.

## THE ANALYTICS REVOLUTION

To understand the potential for government analytics, look no further than the private sector. Tech firms have generated trillions of dollars of value in part or in full by creating, managing, and providing access to diagnostic data. More broadly, the collection and analysis of microdata at previously unforeseen scale has been one of the main drivers of economic and social advancement—from machine-learning algorithms automating customer interactions to utility firms using service data to identify critical weaknesses in physical infrastructure.

Data have enabled firms to improve their own internal operations and management. Productivity is significantly higher among plants that use predictive analytics, an assessment of 30,000 US manufacturer establishments finds (Brynjolfsson, Gorodnichenko, and Kuehn 2021). Units within a business whose workers are engaged have 23 percent higher profit compared to business units with disengaged employees; they also see

significantly lower absenteeism, turnover, and accidents, and higher customer loyalty, the Gallup analytics and advisory organization found, using its database of employee surveys to identify the impact of an organization tracking and nurturing employee engagement (Gallup 2023). In other words, management without measurement—the historically dominant approach to management in firms—puts firms at a competitive disadvantage and undermines productivity. Consequently, an increasing share of private sector companies base their entire business model on analytics (Carrera and Dunleavy 2013).

Sophisticated data collection efforts have uncovered how differences within firms—for instance, across business units—account for the largest share in productivity differences between countries (Cusolito and Maloney 2018). Microdata on firm productivity showcases that variation in firm performance is often due to differences within the firm, such as leadership and management practice (Bloom, Sadun, and Van Reenen 2010; Bloom and Van Reenen 2010). The difference between productive and unproductive firms, it seems, is that unproductive firms allow unproductive units to persist. Or conversely, they do not learn the lessons they could from the most successful units. The most productive firms identify laggards through the use of administrative records or primary data, and target them for improvements or let them go—allowing the firm as a whole to flourish. The private sector, particularly in competitive markets, is disinclined to leave money on the table.

This data revolution in firms *could* be paralleled by one inside government, for at least two reasons. First, government already does analytics on the rest of society. Governments have made significant investments to strengthen the quality of data systems toward better policy making and service delivery, by heavily expanding their measurement of their citizens, firms, and the environment they govern. In most countries, household and firm surveys have become central policy tools to avoid making policy decisions in a data vacuum (Deaton 2003; Kraay 2006). The centrality of such data for state efficacy was striking during the COVID-19 pandemic, for instance, when governments' ability to create effective track-and-trace systems to isolate COVID cases varied widely (Fetzer 2021). In other words, governments have been developing the capabilities for doing analytics on everyone else, but have not turned that lens on their own administrations.

Second, governments also sit on substantial volumes of data that could be used for government analytics. Public administration is frequently symbolized by stacks of government files packed with information on individuals, projects, or tasks. As the use of information and communication technology (ICT) for government operations increases, these records are ever more digitized (Ugale, Zhao, and Fazekas 2019). The digitization of these government records in turn creates a “big data trail,” as a by-product of people's digital behavior. For instance, a microdata point is created every time a firm submits a public procurement bid, a judge makes a court ruling, or a human resource department makes a monthly pay transfer to a public employee. This makes creating the raw material for government analytics far easier than before.

In short, now more than ever, governments are equipped to undertake the government analytics necessary to understand and improve the machinery of public administration. Yet, despite having developed analytic capacity for service delivery by measuring *others* (such as firms and households), many governments are yet to harness the full power of data to measure *themselves* and their own operations. In other words, though government is increasingly developing an evidence base for its services, much further behind is the practice and evidence base for measuring the public administration that actually generates those services.

Existing country-level governance indicators—such as the Worldwide Governance Indicators (Kaufmann, Kraay, and Mastruzzi 2022), the Transparency International (2021) Corruption Perceptions Index, and the Freedom House (2021) Freedom in the World Index—have provided a window into the global distribution of government functioning. However, they are insufficiently granular to inform specific government improvements. Understanding, in broad terms, how effective a government is perceived to be by citizens and firms is helpful, but does not provide guidance on what specific actions governments could undertake to improve effectiveness. Government analytics does just that. Utilizing microdata, it zeroes in on specific inputs, management practices, outputs, and outcomes in specific government organizations, units



within them, and/or public administration as a whole. As such, it puts a more refined sense of reality into each official, manager, and leader's hands. Much as better data to understand what society looks like has transformed public policy for the better, a better sense of what government administration looks like will eventually transform the public administration for the better.

## HOW GOVERNMENT ANALYTICS CAN PROVIDE A STRONGER EVIDENCE BASE FOR IMPROVING GOVERNMENT

Governments across the world make hundreds of thousands of personnel management decisions, undertake millions of procurements, and execute billions of processes each day. The public servants responsible for these activities possess extensive experience and innate knowledge of their respective administrations, which measurement and analytics—no matter how advanced—cannot replace. As one observer noted, “You will never be able to measure away the public sector manager.”

Yet government analytics can provide a stronger evidence base to improve how public officials understand government administration. Rather than substituting for the knowledge of and conversations about the public service, analytics are a strong complement to them. For example, in an experiment with the rollout of a monitoring technology to support agricultural extension workers in Paraguay, researchers found that managers were able to predict which of their staff would benefit most from the program, strengthening its impacts (Dal Bó et al. 2021). As in the private sector and the rest of the economy, analytics and technology are great complements to those who capitalize on them. And without analytics, valuable data trails are being left unexplored, which could lead to missed opportunities for improved decision-making and service delivery.

For example, by utilizing the data sets discussed in the *Handbook*, government officials can recognize the strengths of staff in similar organizations, gain valuable insights to help identify excellent staff in their own organizations, and better allocate their own staff across offices or tasks. This does not mean that managers must abandon their own style; rather, they can learn from the best practices of others. Furthermore, the use of survey data can help governments meet increasing employee expectations for evidence-based staff management, as already practiced by private firms. As in the example that opened this chapter, surveys give employees a voice they would not otherwise have, enriching conversations and providing valuable insights to managers across the service.

Analytics makes internal accountability more effective. When numbers refer to the productivity of an entire sector or a large network of organizations, it can be difficult to hold relevant parties accountable for their performance. However, the microdata embedded in government analytics can enable heads of state to hold the relevant heads of organizations accountable, and heads of organizations in turn can hold the relevant managers of units within the organization accountable. For example, the head of a social security agency can hold directors of regional offices accountable when the speed and quality of social security claims processing at their office falls well below that of other offices. Ultimately, the use of government analytics can enrich conversations in government so that it is more targeted, more engaged with best practices from within the public service, and more grounded in reality.

Government analytics can also enhance the accountability of the government to the public, and improve public conversations about the machinery of government more broadly. By making analytics public, citizens, civil society, and the media can hold the government accountable for how it manages public administration. This can be particularly important when analytics reveal that the machinery of government is not being administered in the public interest. For example, citizens may be interested in knowing how procurement contracts are awarded and whether public sector jobs are based on merit or political grounds. As is the case internally, analytics ensures that public accountability is targeted to organizations where improvements are needed most. That can help avoid unfairly spreading blame across the entire government and reducing the likelihood that any single organization will change (Dunleavy 2017).



## THE OBSTACLES TO ADOPTING GOVERNMENT ANALYTICS

Government analytics thus promises a transformative shift toward evidence-based and continuous improvement of government administration. Why then have many governments not adopted analytics of their administration as proactively as their private sector counterparts?

One reason lies in the lack of systematic evidence on how to do government analytics, and the lack of a systematic compilation of the methods and data available to governments to this end. This *Handbook* is motivated by the need to fill this gap, and hopes to provide a first step toward addressing it.

A second reason is skill shortages for government analytics inside many public sector organizations—both to undertake analytics and to use analytics to improve management. Chapter 3 provides a road map toward addressing these skill shortages, for instance by creating government analytics units inside public sector organizations.

Third, digital records are often primarily designed to enable government operations—such as the award of a contract tender—rather than the analytics of such operations. Governments need to make investments to repurpose government records for analytics—for instance, by connecting, storing, and analyzing data in a cost-effective and secure manner (de Mello and Ter-Minassian 2020). Similarly, online employee surveys are another data source for government analytics. Digitization has made it much cheaper, quicker, and easier for governments to obtain in-depth feedback from employees at scale and to do so more frequently. Again, however, investments in analytics are needed to design, implement, interpret, and make use of employee survey results to improve public administration.

Beyond these obstacles are, however, at least four thornier limits inherent to the analytics of public administration: (1) measurability of outputs and outcomes; (2) institutional complexity; (3) political disincentives to measure; and (4) ethical constraints on measurement. Government analytics requires careful navigation of each of them.

First, not all outputs or outcomes of public administration can be measured. Unlike a private sector firm with a bottom line, public administration organizations have multidimensional missions and tasks whose outcomes and associated processes are often challenging to measure accurately. For instance, how can the quality of policy advice by civil servants to ministers or the quality of a budget a ministry of finance prepares for submission to parliament be measured at scale? Not everything that matters in public administration can be measured. This has made constituencies wary of investing heavily in measurement innovations in the public administration. This is also why analytics templates from private sector firms cannot be indiscriminately copied and applied in public sector organizations.

The response of the *Handbook* to these limits of observability in public administration is to improve, extend, and expand measurement while respecting its limits. This allows public managers to have the best possible knowledge available for administrative improvement. As there are inherent limits to what can be measured in public administration, even with better measurement, public managers need to keep in mind the limits of measurement and triangulate data with other forms of knowledge when devising administrative improvements. Otherwise, as a large literature on performance management in public sectors has found (see, for example, Hood 2006), imperfect measurement can generate a series of unintended and adverse consequences for public management—from gaming performance to effort substitution (expending more effort on measurable metrics of performance at the expense of important but unmeasured outputs and outcomes).

A second challenge is institutional. Public sector reforms struggle with path dependencies (the tendency to become committed to develop in certain ways as a result of structural properties or embedded beliefs and values) (Gains and Stokes 2005). Overhauling data infrastructures and monitoring structures is difficult when organizations fall under different mandates and jurisdictions. Implementation times might extend far into or even span several legislative periods, impairing political incentives for change.

Third, government analytics generates numbers and data that were previously not available. The creation of such data—much like greater transparency in government generally—will sometimes generate political winners and losers (Dargent et al. 2018). To illustrate, creating data on the scale of recruitment

into government based on personal connections rather than merit can generate an avalanche of media reports (Schuster et al. 2020). Government analytics is thus not apolitical. Analysts need to understand what different government actors want to know and what they want to *not* know about the machinery of public administration. Individual analysts and officials will have to negotiate the politics of their setting. Within those constraints, we believe that producing and publishing a broader and more accurate set of government analytics will eventually lead to better conversations within government, and a better government, for society at large.

Last, collecting data on public servants raises ethical concerns, which may limit the scope of analytics. In public administration, ethical considerations are further complicated by trade-offs between individual employee demands for privacy and public demands for values such as productivity, innovation, and accountability of public administrations. Balancing these considerations appropriately is thus central to an ethical pursuit of government analytics, but also implies limits: from an ethics perspective, not everything that can be measured should be measured (see chapter 6).

For understandable reasons, then, the public sector is slower to capitalize on the data revolution than private firms. Much of what it must focus on is harder to measure and requires balancing a greater number of considerations than most private sector activity, whether those considerations are institutional, political, or ethical. Governments have thus fallen behind in their use of modern analytic methods to design and manage public administration. But as the opening stories of how data were successfully used for management improvement in the United States, Brazil, and Nigeria show, this need not be the case. Government analytics can be a powerful tool to improve government if leveraged in the right way.

## GOVERNMENT ANALYTICS AS AN ALTERNATIVE APPROACH TO PUBLIC ADMINISTRATION REFORM

How should governments go about the business of improving government administration? And how should others—such as international development organizations—support these improvement processes? The answer in many countries has been better laws that align governments with “best practices” for the administration of government. However, implementing best-practice public financial management legislation alone has had limited impact in many contexts (Andrews 2013). Similarly, while the number of countries with best-practice merit-based civil service legislation on the books has multiplied, little improvement in merit-based civil service practices has resulted (Schuster 2017). Global best practice may not be an appropriate fit for many country contexts (Grindle 2004, 2007).

Legislation requires a catalyst to make it the reality of government practice. Some observers have urged practitioners to focus on identifying local administrative problems facing particular organizations and attempt solutions in an iterative manner (Andrews 2013; Pritchett, Andrews, and Woolcock 2017). While focusing on addressing specific organizational problems is central to management improvement in public administration, this approach begs immediate questions: What helps officials and governments know which administrative problems are faced by which teams? How can officials know how isolated (or not) those problems are in public sector organizations? And how can we know whether solutions—developed locally or borrowed from other countries or global best practice—have effectively resolved the targeted problems?

This information gap can be addressed at scale through government analytics that leverage governmentwide microdata to diagnose every team in every public organization—for instance, through the data sources outlined in this *Handbook*. This approach can help the government as a whole, specific organizations within it, and individual teams better understand the specific management problems they are facing. Through benchmarking with other teams, organizations, or governments, data analytics can also shed light on opportunities for improvement and who might already be practicing useful approaches. And after governments have undertaken actions to bring about improvements, analytics can help practitioners understand

whether those actions were effective—for instance, in terms of lower procurement costs, lower staff turnover, or better ratings of the quality of managerial leadership according to employee surveys.

Managerial action should, of course, not be taken in isolation based solely on analytics data. As noted, governments will always face measurement challenges and boundaries. Analytical findings must be supplemented with practical and tacit knowledge. Public sector decision-makers thus need to triangulate analytics findings with practical and tacit knowledge. This puts public managers at the heart of interpreting and making use of government analytics findings.

What government analytics can do is strengthen the quality of conversations about how to improve public administration, rather than dictating managerial responses to specific analytics findings. Those conversations about improvement may be in a department of local government, a ministry, or may even span countries and the international community. They may involve government employees, service users, and others, extending beyond managers who are making solitary interpretations of analytics findings. In short, government analytics generates evidence for better conversations—and thus decisions—about how to improve the public administration.

This cycle is, of course, not necessarily linear; it is iterative. At times, government analytics can shed light on management problems that managers were unaware of—as with the opening example of the former marine managing a team in the US federal government. At times, managers may sense a potential problem—as with Brazil’s wage bill—that motivates analytics to better estimate the scale and nature of the problem.

The intention of this *Handbook* is not to dictate a unitary approach to measurement. Instead, the chapters that follow describe many of the most common approaches to government analytics, and present some evidence from across the world for them. This information is offered to provide a framework that government officials and analysts can draw on to select particular analytical approaches of particular use to them. Greater use of government analytics in turn will further the evidence base on how to best undertake government analytics.

This matters because effectively measuring the state allows us to manage it better. The best version of government arises from basing its design and management on the best data achievable as to how government is functioning. Having good policies on the books matters, but much less so without an effective machine to implement them. In other realms, data have revolutionized productivity. It is time to turn the lens on public administration.

## REFERENCES

- Acemoglu, D., and J. A. Robinson. 2019. *The Narrow Corridor: States, Societies, and the Fate of Liberty*. New York: Penguin Press.
- Akhitari, M., D. Moreira, and L. Trucco. 2022. “Political Turnover, Bureaucratic Turnover, and the Quality of Public Services.” *American Economic Review* 112 (2): 442–93.
- Andrews, M. 2013. *The Limits of Institutional Reform in Development: Changing Rules for Realistic Solutions*. New York: Cambridge University Press.
- Besley, T., R. Burgess, A. Khan, and G. Xu. 2022. “Bureaucracy and Development.” *Annual Review of Economics* 14 (1): 397–424.
- Best, M. C., J. Hjort, and D. Szakonyi. 2017. “Individuals and Organizations as Sources of State Effectiveness.” NBER Working Paper 23350, National Bureau of Economic Research, Cambridge, MA.
- Bloom, N., R. Sadun, and J. Van Reenen. 2010. “Recent Advances in the Empirics of Organizational Economics.” *Annual Review of Economics* 2 (1): 105–37.
- Bloom, N., and J. Van Reenen. 2010. “New Approaches to Surveying Organizations.” *American Economic Review* 100 (2): 105–09.
- Bosio, E., and S. Djankov. 2020. “How Large Is Public Procurement?” *Let’s Talk Development*. World Bank blogs, February 5, 2020. <https://blogs.worldbank.org/developmenttalk/how-large-public-procurement>.
- Brynjolfsson, E., Y. Gorodnichenko, and R. Kuehn. 2021. “Predictive Analytics and the Changing Nature of Productivity in Manufacturing.” *Proceedings of the National Academy of Sciences* 118 (13): e2017984118.
- Carrera, L., and P. Dunleavy. 2013. *Growing the Productivity of Government Services*. Cheltenham, UK: Edward Elgar.
- Cusolito, A. P., and W. F. Maloney. 2018. *Productivity Revisited: Shifting Paradigms in Analysis and Policy*. Washington, DC: World Bank.

- Dahlström, C., and V. Lapuente. 2022. "Comparative Bureaucratic Politics." *Annual Review of Political Science* 25 (1): 43–63.
- Dal Bó, E., F. Finan, N. Y. Li, and L. Schechter. 2021. "Information Technology and Government Decentralization: Experimental Evidence from Paraguay." *Econometrica* 89 (2): 677–701.
- Dargent, E., G. Lotta, J. A. Mejía-Guerra, and G. Moncada. 2018. *Who Wants to Know? The Political Economy of Statistical Capacity*. Washington, DC: Inter-American Development Bank.
- Deaton, A. 2003. "Household Surveys, Consumption, and the Measurement of Poverty." *Economic Systems Research* 15 (2): 135–59.
- de Mello, L., and T. Ter-Minassian. 2020. "Digitalisation Challenges and Opportunities for Subnational Governments." OECD Working Papers on Fiscal Federalism 31. Organisation for Economic Co-operation and Development (OECD) Publishing, Paris.
- Dunleavy, P. 2017. "Public Sector Productivity." *OECD Journal on Budgeting* 17 (1): 153–70.
- Fenzia, A. 2022. "Managers and Productivity in the Public Sector." *Econometrica* 90 (3): 1063–84.
- Ferraz, C., F. Finan, and D. B. Moreira. 2012. "Corrupting Learning: Evidence from Missing Federal Education Funds in Brazil." *Journal of Public Economics* 96 (9): 712–26.
- Fetzer, T. 2021. "Did Covid-19 Improve Global Health? A Mixed-Methods Analysis of the Global Health Effect of the Covid-19 Pandemic." *The Lancet* 1 (1): e10–e18.
- Finan, F., B. A. Olken, and R. Pande. 2015. "The Personnel Economics of the Developing State." Chapter 6 in *Handbook of Economic Field Experiments* 2 (2017): 467–514.
- Freedom House. 2021. Freedom in the World. Data file. <https://freedomhouse.org/report-types/freedom-world>.
- Gains, J., and S. Stokes. 2005. "Path Dependencies and Policy Learning in Comparative Public Policy." *Governance* 18 (3): 475–98.
- Gallup. 2023. *State of the Global Workplace: 2022 Report*. <https://www.gallup.com/workplace/349484/state-of-the-global-workplace-2022-report.aspx>.
- Grindle, M. S. 2004. "Good Enough Governance Revisited." *Development Policy Review* 22 (6): 695–703.
- Grindle, M. S. 2007. *Going Local: Decentralization, Democratization, and the Promise of Good Governance*. Princeton, NJ: Princeton University Press.
- Hood, C. 2006. "Gaming in Targetworld: The Targets Approach to Managing British Public Services." *Public Administration Review* 66 (4): 515–21. <http://www.jstor.org/stable/3843937>.
- Kaufmann, D., A. Kraay, and M. Mastruzzi. 2022. Worldwide Governance Indicators. World Bank. Data file. <https://datacatalog.worldbank.org/dataset/worldwide-governance-indicators>.
- Khan, A. Q., A. I. Khwaja, and B. A. Olken. 2019. "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings." *American Economic Review* 109 (1): 237–70.
- Kraay, A. 2006. "When Is Growth Pro-Poor? Evidence from a Panel of Countries." *Journal of Development Economics* 80 (1): 198–227.
- Pepinsky, T. B., J. H. Pierskalla, and A. Sacks. 2017. "Bureaucracy and Service Delivery." *Annual Review of Political Science* 20 (1): 249–68.
- Pritchett, L., M. Andrews, and M. Woolcock. 2017. "Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA)." *World Development* 99: 74–84.
- Rasul, I., and D. Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal* 128 (608): 413–46.
- Schuster, C. 2017. "Legal Reform Need Not Come First: Merit-Based Civil Service Management in Law and Practice." *Public Administration* 95 (3): 571–88.
- Schuster, C., J. Fuenzalida, J. Meyer-Sahling, K. Sass Mikkelsen, and N. Titelman. 2020. *Encuesta Nacional de Funcionarios en Chile: Evidencia para un Servicio Público Más Motivado, Satisfecho, Comprometido y Ético*. <https://www.serviciocivil.cl/wp-content/uploads/2020/01/Encuesta-Nacional-de-Funcionarios-Informe-General-FINAL-15ene2020-1.pdf>.
- Somani, R. 2021. "The Returns to Higher Education and Public Employment." *World Development* 144: 105471.
- Transparency International. 2021. Corruption Perceptions Index. <https://www.transparency.org/en/cpi/2021>.
- Ugale, G., A. Zhao, and M. Fazekas. 2019. *Analytics for Integrity: Data-Driven Approaches for Enhancing Corruption and Fraud Risk Assessment*. OECD (Organisation for Economic Co-operation and Development), Paris.
- Wilson, J. Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.
- World Bank. 2016. *World Development Report 2016: Digital Dividends*. Washington, DC: World Bank.
- World Bank. 2020. Worldwide Bureaucracy Indicators. <https://www.worldbank.org/en/data/interactive/2019/05/21/worldwide-bureaucracy-indicators-dashboard>.
- World Bank. 2021. *World Development Report 2021: Data for Better Lives*. Washington, DC: World Bank.
- World Bank. 2022. "GovTech Maturity Index, 2022 Update: Trends in Public Sector Digital Transformation." Equitable Growth, Finance and Institutions Insight–Governance. World Bank, Washington, DC.
- World Bank Group. 2019. *Innovating Bureaucracy for a More Capable Government*. Washington, DC: World Bank Group.

## CHAPTER 2

# How to Do Government Analytics

## Lessons from the *Handbook*

*Daniel Rogger and Christian Schuster*

### SUMMARY

How can practitioners and researchers undertake government analytics effectively? This chapter summarizes lessons from *The Government Analytics Handbook*. The chapter begins by introducing a public administration production function, which illustrates how different data sources, such as procurement data or public servant survey data, shed light on different parts of the machinery of government. The chapter then highlights lessons to keep in mind when undertaking any measurement and analysis of government administration. In the chapter-by-chapter summary that follows, lessons on how to generate and analyze a range of core data sources for government analytics are presented. These are data from administrative records (such as human resources and payroll, budget, procurement, administrative cases, task completion and projection completion); data from surveys of public servants; and external assessments of government, such as household or citizen surveys. The chapter concludes by showcasing how different data sources can be integrated to understand core challenges in the administration of government, such as personnel management.

### ANALYTICS IN PRACTICE

- This *Handbook* presents a wealth of approaches and data sources available to governments to improve the analytics of the machinery of government and identify evidence-based improvements. Many of these approaches rely on data that governments already collect as part of their day-to-day operations.
- By conceiving of government's electronic records as data in themselves, existing government records, and in particular the vast troves of data now being produced, can be repurposed as a means of diagnosing and strengthening government administration.
- Government analytics can be undertaken with at least three types of data: administrative data collected or published by government entities (such as payroll data); surveys of public servants; and external assessments (such as household surveys or anthropological assessments).



- To organize the various data sources assessed throughout the *Handbook*, this chapter introduces a public administration production function. A production function relates input factors of production through processes (such as management practices) to the output of an organization, and their eventual outcomes.
- Some data sources are better suited to assessing inputs into public administration, such as payroll data assessing the costs of different personnel. Some data sources are better suited to assessing the processes, practices, and cultures that convert inputs into outputs, such as surveys of public servants that assess how they are being managed. And some data sources are better suited to assessing the outputs and outcomes of public administration, such as citizen satisfaction surveys. What type of data source is appropriate for analytics depends on what aspect of public administration the analyst is seeking to diagnose and improve. Data sources can also be powerfully combined to understand overarching themes in how government is functioning, such as corruption.
- Frontier government analytics would integrate the analytics of these distinct data sources. It would generate them at a scale sufficient to inform the decision-making of individual managers. And it would make them easily accessible to those managers across government organizations and departments. For instance, dashboards integrating data sources and updating in real time would provide managers with comparisons for their staffing issues, process quality, the perceived quality of management practices, and so on. They could track outputs and outcomes, from task completion and case productivity to external assessments from citizens. Comparative data would allow them to benchmark themselves against other government organizations, or where appropriate, other countries. Managers would be capable of understanding the limitations and strengths of different analytics. The result would be a transformational change toward leveraging data to strengthen public administration.

## INTRODUCTION: AN UNPRECEDENTED EXPLOSION OF DATA

A common perception of government is that it is a place that no one truly understands: an incomprehensible maze, filled with a procedural fog. Actually, public administration is brimming with information. The everyday business of government has generated an abundance of records on who does what where. Each of these records has its origin in a contemporary administrative need of public officials. With the right approach, these records also provide a window into the workings of the administrations that created them—a way to clear the fog and provide a map through the maze.

For example, as a social security officer decides how to organize their work, they are making decisions that will affect the delivery of cases they are working on. These choices are reflected in the time stamps that accompany their cases, the degree to which they follow procedure, and the feedback they receive from those they serve. On its own, the record of their work is interesting, but combined with the records of their colleagues from their own organization and across the country it becomes a means to understanding the functioning of a public sector organization. Do all the officers working for a particular manager perform well on some aspects of their work but less so on others? Are some offices able to regularly process retirement benefits faster than others? The fog of uncertainty about how social security officers perform across their community can be cleared by turning records into analytics of the public service.

The data trail of government records like these has expanded greatly recently. The move toward digital government has multiplied the scale of electronic transactions between public administrations and citizens. From citizens paying taxes electronically to registering online for driver's licenses, every interaction is now recorded electronically in government data systems. Digital government has also multiplied electronic transactions within public administrations. Payrolls, procurement records, and budget disbursement records are some examples. Digital government has also facilitated the collection of other data, for instance by enabling governments to survey their employees at scale online. All this makes data to understand public administration easier to access than ever before.

This trove of data is critically important to the productive functioning of government, much in the same way as it has improved firm performance (Bakshi, Bravo-Biosca, and Mateos-Garcia 2014). “Without high-quality data providing the right information on the right things at the right time, designing, monitoring and evaluating effective policies becomes almost impossible,” a United Nations report notes (UN 2014, 2).

While the use of data for public *policy* has exploded, far less emphasis has been placed on how government records might be repurposed to understand how effectively the administrative machinery of government is functioning. By conceiving of government’s electronic records as data in themselves, existing government records can be seen as a means of diagnosing and strengthening government administration.

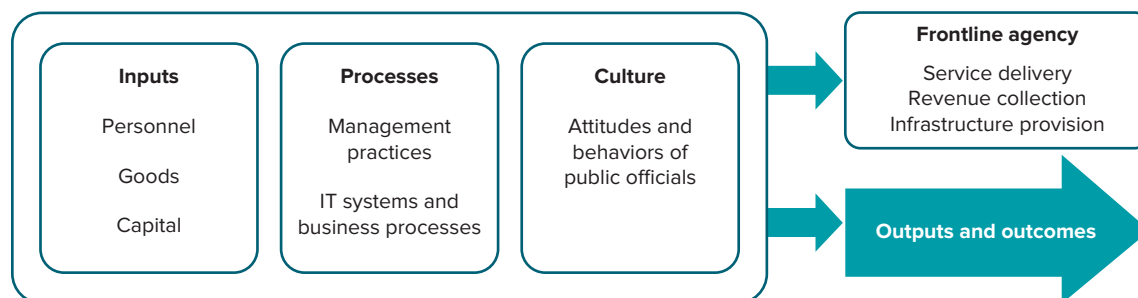
As discussed in chapter 1, public administration presents particular measurement challenges and opportunities. This *Handbook* focuses on sources and methods of measurement to meet those challenges, and on the need to create a system of measurements that is fit for public administration, rather than repurposed from other settings. Although current government records are an excellent foundation for analysis, their suitability for generating knowledge requires work. Once analytics becomes a goal, the range and nature of the public record is likely to change, to optimize government data for both the task at hand and the broader assessment of government functioning. This chapter summarizes lessons to that end—and on how to do government analytics—from across the *Handbook*.

## WHAT GOVERNMENT ANALYTICS CAN ANALYZE: UNDERTAKING ANALYTICS ALONG A PUBLIC ADMINISTRATION PRODUCTION FUNCTION

Government analytics refers to the use of data to diagnose and improve the machinery of government, or public administration. This chapter introduces a public administration production function to provide an overarching framework to organize the various data sources assessed in different chapters.<sup>1</sup> A production function relates input factors of production to the output of deliverables of an organization, and their eventual outcomes. The productivity of an organization thus depends on the quality and quantity of outputs relative to inputs. Figure 2.1 visualizes the different components of our production function for public administration (Meyer-Sahling et al. 2021; World Bank Group 2019). While many core elements coincide with typical private sector production functions (Mas-Colell, Whinston, and Green 1995), the functioning of government administration has been characterized as distinct from that of private firms due to the multiplicity of principals, the ambiguity of tasks, and the presence of principals with political incentives, among other features.

In public administration, inputs include personnel (public employees), goods (such as computers), and capital (such as office space). Outputs refer to, first, the deliverables produced by public

**FIGURE 2.1** The Public Administration Production Function



Source: Adapted from World Bank Group 2019.  
 Note: IT = information technology.



administration organizations themselves. For instance, a ministry of finance might issue public sector debt at a certain interest rate. Further, public administration organizations produce outputs (activities) that enable frontline agencies in the public sector—such as hospitals, schools, or police forces—to deliver services and goods to citizens. The outcomes in these examples are better health, education, or public safety, respectively. To fund the outputs, a ministry of finance may oversee budgets that frontline agencies then disburse to deliver their services.

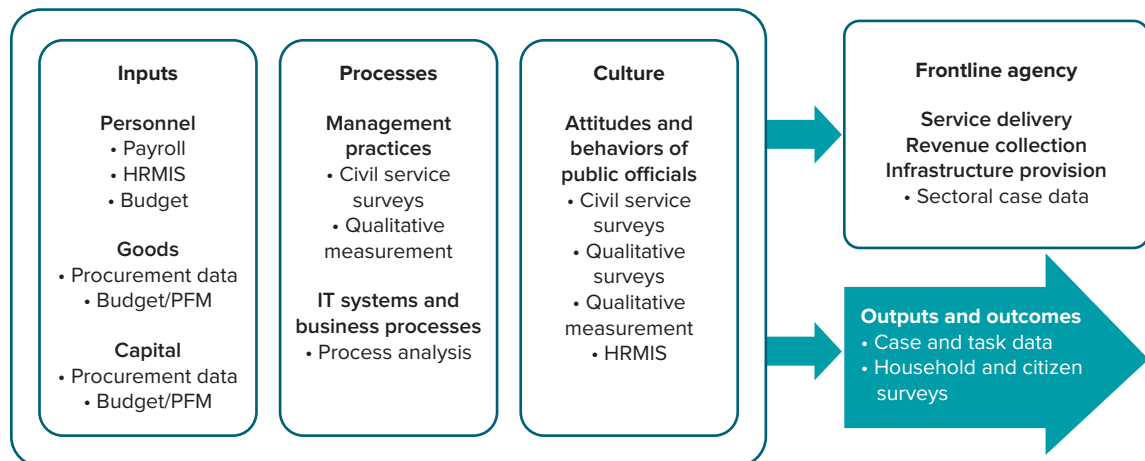
How do public administrations convert inputs (such as personnel) into outputs and outcomes? In our production function, this conversion is enabled by policies (organizational objectives and work procedures), systems, and management practices, and mediated by norms and behaviors inside public administration. For instance, a ministry of finance may have a policy in place to review a budget for an organization by a certain date. A team lead inside the ministry then manages employees to ensure the task is completed well and on time—such as through effective performance management practices. Those practices and organizational policies shape the norms and behaviors of the ministry’s employees—such as their motivation to work hard—which in turn then allows the ministry to produce outputs (such as a budget review).<sup>2</sup>

By utilizing different data sources and different methods, government analytics can shed light on all parts of this production function and identify bottlenecks, whether overpriced input goods, ghost workers on the payroll, high staff turnover, or slow processing of administrative cases, to name just a few. Contemplating government analytics along the production function enables analysts to diagnose public administration challenges holistically, and to understand how different data and approaches to government analytics relate.

To illustrate, figure 2.2 maps a number of different data sources analyzed in various chapters to their respective components in the production function. Several types of administrative data have particular strengths in diagnosing inputs into the public administration production function. For instance, payroll data and human resources management information system (HRMIS) data can help governments understand personnel as an input into the public administration production function, such as whether pay of public servants is competitive and fiscally sustainable, or whether staffing levels are adequate (see chapters 9 and 10). Budget data and procurement data can help governments understand spending on goods and capital as inputs into public administration—for instance, whether public administrations acquire similar goods cost-effectively across organizations in the public administration (see chapters 11 and 12).

Government analytics can also shed light on the processes and practices that convert inputs into outputs and outcomes. Surveys of public servants and qualitative measurement have particular strengths

**FIGURE 2.2 Mapping Different Government Analytics Data in the Public Administration Production Function**



Source: Original figure for this publication adapted from World Bank Group 2019.

Note: HRMIS = human resources management information system; IT = information technology; PFM = public financial management.

at diagnosing management practices. Management quality is fundamentally experienced by employees and a result from the interaction between managers and employees. Surveys can, for instance, ask public servants how they perceive the leadership of their superior or the quality of their performance feedback (see chapter 18). Government analytics can also shed light on the quality of processes inside public administration, such as whether these processes adhere to government procedure or meet deadlines (see chapter 13).

Whether practices and processes effectively turn inputs into outputs and outcomes is, as noted, mediated by the norms, attitudes, and behaviors of public administrators. Surveys of public servants and qualitative measurement are standard practice in many governments to evaluate this component of public administration production—for instance, to understand how engaged, committed, and ethical in their behavior public administrators are (see, for example, chapter 18). HRMIS data often complement rich survey data by providing insights into specific behaviors of public employees that are digitally recorded, such as whether public servants leave the organization, work overtime, or take sick leave (see chapter 9).

Last, public administrations produce outputs and outcomes both of their own (such as a ministry of finance issuing debt), and to enable outputs and outcomes of frontline providers. The productivity of frontline, service delivery agencies such as hospitals, schools, and police forces has been extensively measured, not least as direct contact with citizens enables more direct measurement of service delivery outcomes (such as patient outcomes in hospitals or learning outcomes in schools) (see chapter 29).

This *Handbook*, instead, focuses on the analytics of *administrative* outputs and outcomes. Administrative case data are one important source for measurement in such contexts. Such data are often routinely collected by organizations (for instance, the number of tax or social security cases processed) and can be repurposed by organizations to measure outputs and outcomes (such as the amount of tax revenue raised), and thus gauge productivity (see chapters 14 and 15). Beyond administrative data, surveying households and citizens (such as by asking citizens about their trust in public administration organizations) can be an important data source to understand outcomes of public administration (see chapter 28). What will be of most significance to measure and analyze will depend on a specific organizational setting and topic of interest to decision-makers.

The various chapters of this *Handbook* provide insights into how to use these different data sources to understand how government is functioning and improve the management of public administration. For a brief overview, table 2A.1 in annex 2A at the end of this chapter presents a chapter-by-chapter mapping of data sources to the topics covered in chapters 10 to 30 of the *Handbook*, to help readers pick and choose the chapters most of interest to them. The remainder of this chapter thus summarizes their lessons on how to do government analytics well.

## PART 2: FOUNDATIONAL THEMES IN GOVERNMENT ANALYTICS

The second part of the *Handbook* focuses on four cross-cutting challenges in the measurement and analysis of public administration: how to ensure that government analytics are undertaken (1) with due recognition and management of the risks involved (from managing political pressure to protecting key indicators from manipulation); (2) in a way that ensures analytical practices accord with current best practice in social science; (3) in an ethical manner; and (4) with measures to determine whether the government analytics generated are, in fact, used.

### Measuring What Matters: Principles for a Balanced Data Suite That Prioritizes Problem Solving and Learning

Chapter 4, by Kate Bridges and Michael Woolcock, tackles the first of these four challenges. The chapter notes that overreliance on quantitative government analytics comes with risks. For example, as with any performance targeting, hitting easy-to-measure targets risks becoming a false standard of broader success. Overemphasis on measurement also risks neglect of other important forms of knowledge—such as

qualitative work or practical knowledge—and thus may curtail a deeper understanding of key public administration problems. Moreover, political pressures, if undetected, can lead to falsification of data—for instance, to cover up problems such as corruption—and thus undermine the integrity of government analytics. Left unattended, these risks may mean that government analytics may curtail the problem-solving and implementation capabilities of public sector organizations, rather than strengthening them.

In light of these risks, chapter 4 offers four principles for government analytics based on a balanced data suite that strengthens the diagnosis and solving of problems in public administration:

1. Identify and manage the organizational capacity and power relations that shape data management: for instance, by defining and communicating to all staff professional standards for collecting, curating, analyzing, and interpreting government analytics data.
2. Focus quantitative measures of success on those aspects that are close to the problem: for instance, by targeting measurement of public administration problems prioritized by government.
3. Embrace a role for qualitative data, especially for those aspects that require in-depth, context-specific knowledge.
4. Protect space for judgment, discretion, and deliberation in those (many) decision-making domains that inherently cannot be quantified.

### **Practical Tools for Effective Measurement and Analytics**

Chapter 5, by Maria Ruth Jones and Benjamin Daniels, turns to a second cross-cutting challenge: how to employ the best practices of modern social science when utilizing statistical tools for government analytics. The chapter discusses important underlying statistical principles for government analytics to ensure that government analytics are credible, including the transparency of analysis and reproducibility of results. Producing analysis that accords with best practice requires considering the full life cycle of data work, such that each stage of handling data can be designed to support the next stages. The chapter introduces the suite of tools and resources made available for free by the World Bank's Development Impact Evaluation (DIME) Department to support the achievement of best-practice statistical analysis, such as research-cycle frameworks, extensive training tools, detailed archives of process and technical guidance, and a collaborative approach to data and analytics. The chapter also links to online tools, such as an online training hub, available to help implement these practices.

### **The Ethics of Measuring Public Administration**

Chapter 6, by Annabelle Wittels, discusses a third cross-cutting challenge in government analytics: how to undertake ethical measurement and analysis in public administration. While guides for the ethical collection of data on citizens exist, there is a dearth of practical guides on the ethics of government analytics, particularly with respect to data collection by governments on their own employees. Chapter 6 introduces a heuristic to encourage ethical government analytics, which balances three, at times competing, ethical considerations: (1) an individual dimension that encompasses demands by public employees for dignity and privacy; (2) a group dimension that relates to allowing for voice and dissent of public employees; and (3) a public-facing dimension that ensures that analytics enable public administrators to deliver on public sector values—accountability, productivity, and innovation. These three considerations can be in tension. For instance, data diagnostics on public personnel can inform better management practices and enable greater productivity but impinge on the privacy of employees, whose data are required for the diagnostic.

To guide practitioners, chapter 6 presents a 10-point framework. For instance, the ethics of government analytics requires consideration of the granularity of data required for the diagnostic. Are data that identify individuals required for the diagnostic, or could group-level or anonymous individual-level data enable a

similar diagnosis? In a survey of public servants, do demographic questions need to identify the exact age of a respondent in an organization (which risks identifying the respondent), or are broad age bands (with a lower risk of identification) sufficient? As a second example, ethical government analytics requires consideration of who has a say in what gets measured. In particular, are those who are being measured consulted in advance and given an opportunity to provide inputs? Questions in the framework like these can guide ethical government analytics.

### Measuring and Encouraging Performance Information Use in Government

Chapter 7, by Donald Moynihan, discusses efforts to address a fourth challenge in government analytics: measuring whether government analytics measures are, in fact, being used. If costly data collection or analysis is undertaken on public administration but the resulting analytics are not used to improve public administration, government analytics harms rather than advances better government. Chapter 7 illustrates how survey and administrative data can be drawn on to measure use of government analytics data (or, to use the term in chapter 7, performance information). In particular, the chapter recounts the experience of the US Government Accountability Office (GAO), which periodically surveys public employees on their use of performance information. The chapter describes the survey measures deployed to this end. The chapter also discusses administrative data that governments can utilize, such as data that track the use of dashboards displaying government analytics data. The use of government analytics can thus be subjected to the same evaluative measurement rigor that government analytics applies to public administration.

### Understanding Corruption through Government Analytics

Chapter 8, by James Anderson, David S. Bernstein, Galileu Kim, Francesca Recanatini, and Christian Schuster, illustrates how the approaches presented in this *Handbook* can be combined and leveraged to holistically diagnose a major obstacle to more effective public administrations. Chapter 8 is discussed further at the end of this chapter.



## PART 3: GOVERNMENT ANALYTICS USING ADMINISTRATIVE DATA

The nine chapters in part 3 discuss the range of administrative data sources that governments can draw on to undertake government analytics. Each main type of administrative data is covered in a different chapter. Chapters 9 and 14 contextualize these discussions, showcasing how to create underlying data infrastructures (chapter 9) and how to combine data sources to measure the performance of public organizations with multidimensional missions (chapter 14).

### Creating Data Infrastructures for Government Analytics

Chapter 9, by Khuram Farooq and Galileu Kim, focuses on how to create data infrastructures—or management information systems (MIS)—that are well suited for government analytics. Using the case of human resources management information systems (HRMIS), the chapter provides a road map to guide the development of data infrastructures that enable government analytics. The road map emphasizes the importance of first getting in place high-quality foundational data modules in the MIS, and only then transitioning to developing more complex analytical modules. In the case of HRMIS, for instance, ensuring high-quality foundational modules including basic information on personnel and payroll compliance should take precedence over more advanced modules such as talent management and analytics. Without quality foundations,

other modules will produce imprecise or inaccurate analytics. Analytical modules that include dashboards and reports require that foundational modules are set in place and their data are accurate.

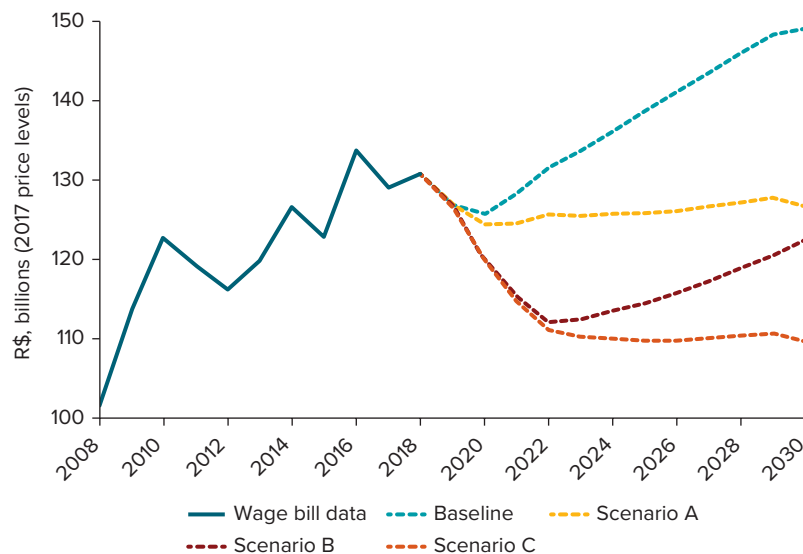
The road map thus emphasizes a sequential approach to creating data infrastructures, coupled with repeated testing of new infrastructures, accessible technical support for users of new MIS systems, tracking of usage, and building in-house capacity to maintain the system. To illustrate potential applications, the chapter is complemented by three case studies of analytical transformations in HRMIS systems in Luxembourg, Brazil, and the United States. Ludwig Balmer, Marc Blau, and Danielle Bossaert discuss how the government of Luxembourg introduced a Business Intelligence Center for human resources (HR) analytics, which transformed HR operations. Luciana Andrade, Galileu Kim, and Matheus Soldi Hardt showcase the development of a system to detect irregularities in Brazil’s federal payroll, using machine learning. Robin Klevins and Camille Hoover illustrate how a US federal government agency developed a simple but effective dashboard to extract insights from a federal public employee engagement survey.

The quality of analytics arising from any MIS is critically dependant on the quality of the underlying data, and the analysts’ understanding of their origins and limitations. The rest of part 3 discusses how to ensure quality measurement and analytics with a range of administrative data sources that measure aspects of the production function of public administration.

### Government Analytics Using Human Resources and Payroll Data

In chapter 10, Rafael Alves de Albuquerque Tavares, Daniel Ortega Nieto, and Eleanor Florence Woodhouse illustrate how to use payroll and HRMIS data for government analytics. Using a series of examples from Latin American countries, the chapter underscores how the analytics of such data can enable more evidence-based decisions around both fiscal planning and personnel policy. For instance, payroll data can be drawn on to better understand the likely future dynamics of wages and retirement based on the modeling of individual career trajectories. Figure 2.3 illustrates how analytical use of payroll data allowed Brazilian policy makers to simulate the financial implications of a set of policies related to pay and employment. The difference between the wage bill costs under the least and most expensive options amounted to nearly 50 percent

**FIGURE 2.3 Wage Bill Projection and Policy Scenarios, Brazil, 2008–30**



Source: Original figure for this publication (see chapter 10).

of total wage bill expenditures. Such a granular lens on employment trends enables governments to better plan for the longer-term fiscal implications of public employment and understand its drivers.

The analytics of payroll and HRMIS data can also enable governments to improve personnel policy. For instance, data on workforce allocation across different regions of a country (such as those of a tax administration) matched to data on the number of service users (such as taxpayers) in each region can help governments understand under- or overstaffing and improve workforce allocation by prioritizing new recruitment in understaffed regional offices. Payroll data also enable governments to measure turnover of employees working in different organizations or for different managers, helping governments pinpoint retention problems and their origins. By comparing pay for similar positions across government institutions, payroll analysis identifies potential salary inequities. Chapter 10 recommends centralizing payroll and HR data collection systems to render such data accessible and provides a road map to this end that complements the discussions and case studies in chapter 9.

### **Government Analytics Using Expenditure Data**

Chapter 11, by Moritz Piatti-Fünfkirchen, James Brumby, and Ali Hashim, discusses government analytics using expenditure data. Budget and government expenditure data are already widely used by governments to understand whether resources are used for budgeted priorities; whether spending is sustainable, efficient, effective, and equitable; and which government transactions (typically, large-value ones) might have high fiduciary risks. The World Bank, for instance, has a long-standing practice of Public Expenditure Reviews. Such reviews are often accompanied by benefit incidence analysis to orient spending toward those in need, by linking data on spending distribution with who receive services. The chapter briefly reviews these well-established uses of expenditure government analytics, and then delves into an aspect missing in much expenditure analytics: how to ensure high-quality data for expenditure analytics.

The chapter underscores the need to periodically review government expenditure microdata for five features: (1) data provenance and integrity; (2) comprehensiveness; (3) usefulness; (4) consistency; and (5) stability. This requires a prior, clear definition of what counts as a government expenditure, as well as a clear understanding and documentation of how transactions across spending items in government are created, what control protocols they are subject to, how this information is stored, and how microdata are aggregated for analysis (such as by classifying government transactions by function, to be able to analyze aggregate spending data by function). The chapter provides practitioners with seven questions to include in expenditure analytics to ensure that the data are high quality, and the resulting analytics are as informative as possible to improve government spending.

### **Government Analytics Using Procurement Data**

Chapter 12, by Serena Cociolo, Sushmita Samaddar, and Mihaly Fazekas, discusses how to use procurement records as data for government analytics. The digitization of national public procurement systems across the world has multiplied opportunities for procurement data analytics. Such analytics allow governments to strategically monitor procurement markets and trends, to improve the procurement and contracting process through data-driven policy making—for instance, by identifying overpriced goods or corruption risks in procurement—and to assess the potential trade-offs of distinct procurement strategies or reforms. The chapter explores the range of procurement indicators that can serve these purposes. For instance, indicators to measure the economy and efficiency of procurement include the time needed for contracting and the final price paid. Indicators to proxy transparency and integrity include the share of single bidder tenders and the share of excluded bids. Indicators to measure competition include the number of bidders. Indicators of inclusiveness and sustainability include the share of bids coming from small and medium enterprises.

When e-procurement systems are integrated with other e-government systems—such as systems generating firm registries or tax data—analytics can go even further: for instance, by allowing governments



to detect potential family relations (and thus collusion risk) between owners of firms bidding for government contracts and procurement officials. Chapter 12 also showcases how governments can use interactive dashboards to track, analyze, and display key procurement indicators through customizable and user-friendly visualizations. All this requires that procuring entities record procurement transactions consistently, that such records are then centralized, and that (automated) data quality checks and periodic data audits ensure the data are accurate.

### Government Analytics Using Data on the Quality of Administrative Processes

The eventual value of the core inputs to the effective functioning of government (personnel, goods, and capital) is determined by how they are processed and managed. Chapter 13—by Jane Adjabeng, Eugenia Adomako-Gyasi, Moses Akrofi, Maxwell Ampofo, Margherita Fornasari, Ignatius Geegbae, Allan Kasapa, Jennifer Ljungqvist, Wilson Metronao Amevor, Felix Nyarko Ampong, Josiah Okyere Gyimah, Daniel Rogger, Nicholas Sampah, and Martin Williams—presents approaches to assessing the quality of *administrative* processes. Applying proper processes and procedure to a project, file, or case is core to the work of public administrators. The chapter presents a range of indicators to this end, such as the share of processes undertaken by public administrators that are timely with respect to deadlines, adhere to government procedure, and are logical in flow. Such data can be collected automatically as part of digitized government work, or manually by assessors employed to judge the quality of process in the physical records of projects, files, or cases. Chapter 13 showcases two applications of this approach. The example from Liberia highlights adherence to new processes for performance appraisal. The example from Ghana highlights the quality of process in core office duties, such as project planning, budgeting, and monitoring.

### Government Analytics Using Customs Data

Chapter 14, by Alice Duhaut, provides an overview of government analytics using customs data. Customs agencies typically have three core objectives: facilitating trade, collecting revenue, and ensuring the security and safety of the goods entering or exiting the country. As in many other government agencies with multidimensional missions, attaining one objective (such as greater safety of traded goods) can come at the expense of another (such as facilitating trade). Incomplete measurement of objectives risks encouraging attainment of measured objectives while unknowingly impairing other objectives. This puts a premium on effective measurement of all dimensions of a customs mission, which often requires triangulating different data sources. Chapter 14 showcases how this can be done, deriving indicators for trade facilitation (such as costs of the process, particularly in terms of delays); revenue collection (such as trade volume and revenue collected based on the assessed value); and safety (such as number of goods in infraction seized). Collecting these indicators requires integrating multiple data sources. The chapter thus discusses several data sources. These include the Automated System for Customs Data, used by 100 countries and designed by the United Nations Conference on Trade and Development (UNCTAD), which captures items declared, excise, and duties; as well as complementary data sources, such as time-release studies and GPS data on the time spent at borders. The chapter also illustrates how such data can be used not only for risk management *ex ante*, but also to assess customs performance *ex post*.

### Government Analytics Using Administrative Case Data

Chapter 15, by Michael Carlos Best, Alessandra Fenizia, and Adnan Qadir Khan, provides insights into the analytics of administrative case data in government more broadly. A case file is typically a collection of records regarding, for instance, an application for government payments (such as social security) or access to services (such as government-sponsored child care); to obtain licenses and permits; or to bid on a government contract. The chapter draws on three example types of administrative cases: social security programs,



tax collection, and public procurement. In all three examples, governments routinely collect case data as part of their day-to-day operations. The chapter shows how these case data can be repurposed to construct objective measures of performance. These include measures to benchmark productivity across, say, regional tax offices, such as the average time to complete similar tax cases or the number of cases completed per officer. Map 2.1 presents an index of the productivity of social security case processing in Italy. It illustrates how substantial the variability in government productivity can be across a single country, with some offices taking 2.5 times as long as others to process similar cases. Case data also enable governments to understand differences in quality, such as comparing the share of tax cases leading to appeals by taxpayers or being identified as erroneous by audits.

Chapter 15 also emphasizes the importance of accounting for the complexity of each case. For example, a social security claim that clearly meets the requirements of regulation is less complicated to process than a case in which there are ambiguities in eligibility and external validation is required. The chapter provides guidance on how to adjust case data for complexity. When accounting for complexity and quality, the chapter concludes that the case data governments already collect provide a wealth of performance information to make such adjustments.

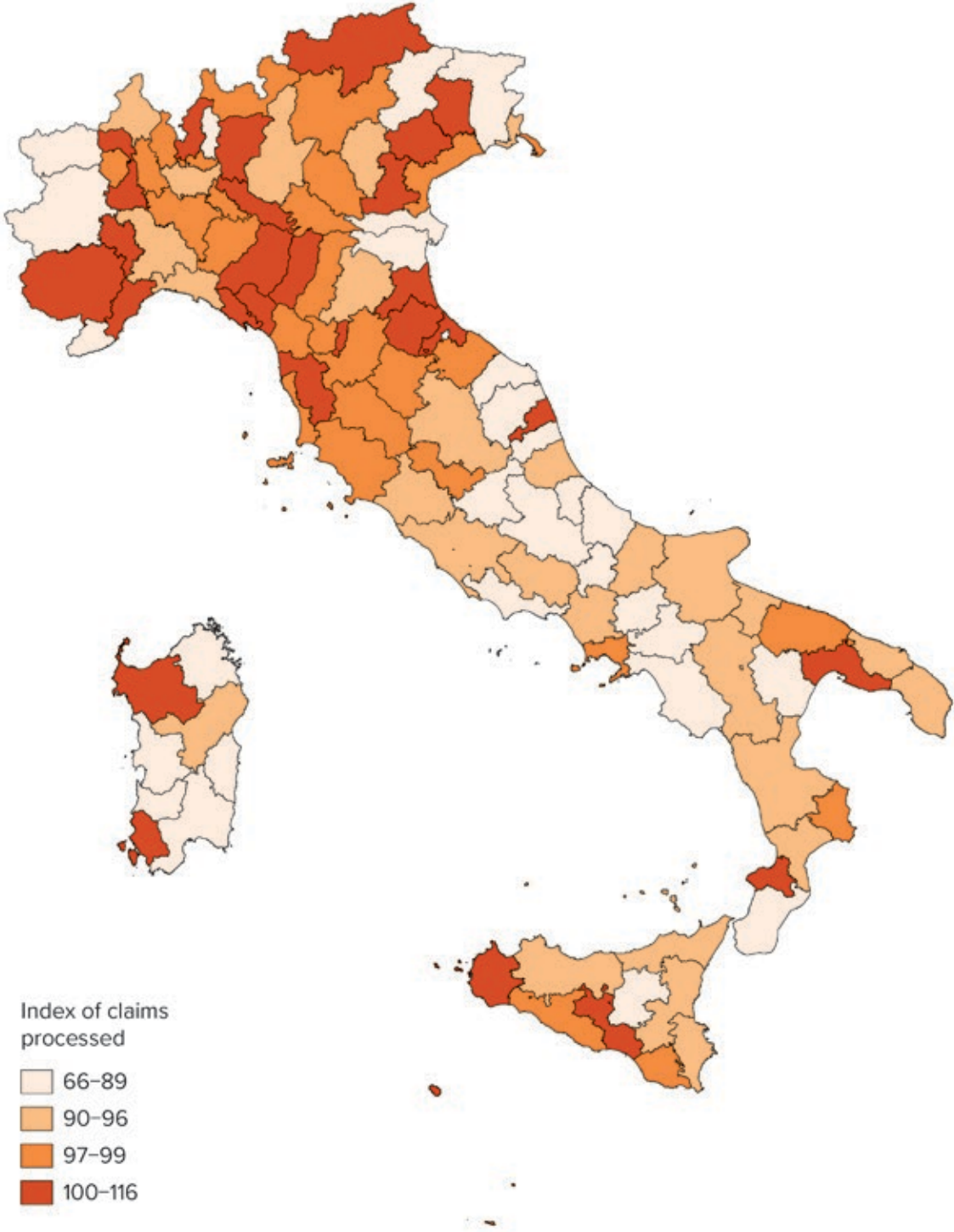
### Government Analytics Using Machine Learning

Chapter 16, by Sandeep Bhupatiraju, Daniel Chen, Slava Jankin, Galileu Kim, Maximilian Kupi, and Manuel Ramos Maqueda, shifts the focus to a different data source—text-as-data—and a different methodological approach, the use of machine learning (ML) and artificial intelligence (AI) for government analytics. Machine learning is fundamentally a methodological approach: it defines a performance indicator and trains an algorithm to improve this indicator, using the data collected. Such data can include text, allowing government to classify quantitatively the “big data” of texts it produces in records or communications. As a result, ML can be applied in a range of government analytics domains, from detecting payroll fraud to understanding bias in welfare appeal decisions, to name a few. The chapter illustrates the use of ML and AI for government analytics in the case of the judiciary. In the justice system, the increasing digitization of legal documents and court sentences, and the development of new techniques in natural language processing, enable analytics to improve judicial decision-making. India, for instance, has 27 million pending court cases; the sheer number of cases precludes manual classification of (often inconsistent) records and legal texts. ML algorithms can be trained to classify such records. This enables, for instance, analytics of bias and discrimination in courts (such as where judges with certain characteristics are associated with certain judicial outcomes in similar cases), or evaluations of how judicial reforms shape judicial productivity and bias. The chapter also describes the enabling environment for ML application—including how to build ML human capital and data infrastructure; the ethical considerations to keep in mind; and the importance of collaboration between ML engineers, domain experts, and the agencies that will use the technology to develop effective ML-based government analytics.

### Government Analytics Using Data on Task and Project Completion

Chapter 17, by Imran Rasul, Daniel Rogger, Martin Williams, and Eleanor Florence Woodhouse, discusses government analytics using task completion data. Much government work consists of the completion of tasks, from creating major reports to undertaking training programs and building infrastructure. A task completion approach allows for the investigation of which units and organizations are most likely to initiate, make progress on, and complete tasks—particularly where organizations complete similar tasks (such as preparing budgets). A task completion approach is particularly important to understand performance in administrative organizations in which the productivity data discussed in previous chapters (such as case data or frontline service delivery indicators) are not available.

**MAP 2.1** Variations in Productivity of Processing Social Security Cases, Subregions of Italy



Source: Fenizia 2022, using Italian Social Security Agency data.  
Note: The key refers to the number of social security claims of a particular type that are processed by an office in a particular time period divided by the full-time equivalent of workers of that office during that time.

Data for task completion can be extracted from a variety of sources. In Ghana, for instance, government units complete quarterly progress reports on all activities, which can be digitized and repurposed into task completion data. In the United Kingdom, the National Audit Office completes major project reports that assess the progress of large infrastructure projects against corresponding planning documents. By subsetting—that is, culling—these data to similar tasks undertaken by all government units, units can be benchmarked on the average time it takes them to complete tasks and the share of uncompleted tasks, for instance. Matching these data to other data about these units—for instance, on budget disbursements or management practices—can help governments understand why some government units are more effective at completing tasks and identify drivers to improve productivity in lagging government units. Chapter 17 underscores the importance of—and provides guidance for—classifying task characteristics correctly (such as in terms of their complexity) to ensure that cross-unit benchmarking is valid.

### Cross-Cutting Insights from Part 3

The chapters in part 3 reveal some cross-cutting insights to keep in mind when undertaking government analytics using administrative data. To begin with, high-quality administrative data are not a foregone conclusion. The infrastructure for—and thus quality of—the underlying measurement and resulting data is paramount and requires careful consideration. The first challenge is data coverage. A central payroll system, for instance, might cover only part of the public employment workforce (chapter 10); and a central financial management information system might cover only part of government expenditures (chapter 11). A second challenge is data validity. To illustrate, case data in tax or social security are often not recorded or designed to measure performance. Careful thought is needed to repurpose such data for performance measurement. Self-reported data—as in the case of some task completion data—may also suffer from inaccurate or manipulated data entry by the units being evaluated, putting a premium on independent, third-party data collection or validation (chapter 7). A third challenge for performance data in particular is completeness relative to the mission of an organization. As chapter 14 notes, missions and goals of public sector organizations are typically multidimensional (and at times contradictory). This requires the triangulation of multiple data sources to ensure performance is measured holistically, so that government analytics does not incentivize the attainment of one goal of an organization at the expense of another (chapter 4).

The chapters also emphasize the human capital requirements for government analytics using administrative data. Creating the information technology (IT) systems that underlie centralized data infrastructures requires IT and data science skills (chapters 9 and 16). Processing records into usable data, analyzing the data, and making the data accessible for analysis and management improvements (such as through dashboards) similarly require data science and visualization skills. In some governments, advanced data science skills (such as for machine learning) might be in short supply. Upskilling data scientists in government, or creating data science teams for government analytics, is thus important to make the most of government analytics opportunities.

Lastly, the chapters also emphasize the importance of data accessibility for decision-makers. Creating dashboards to allow users to explore key insights from the data—for instance, on procurement indicators or HR—facilitates such accessibility. These can be complemented by briefings and slide decks with key diagnostic findings and insights and data storytelling for senior policy makers to take action. In other words, government analytics data do not speak for themselves, but need to be made understandable to support government action.

## PART 4: GOVERNMENT ANALYTICS USING PUBLIC SERVANT SURVEYS

Part 4 focuses on a single data source for government analytics: surveys of public servants. There are three reasons for dedicating an entire section to one data source.

First, surveys of public servants are one of the most widely used data sources for government analytics. The review of the global landscape in chapter 18 finds that the number of governments implementing governmentwide surveys of public servants repeatedly every year or two has increased continuously over the last two decades. At least nine governments of member countries of the Organisation for Economic Co-operation and Development (OECD) were conducting annual or biannual surveys as of 2021, and many others are surveying their public servants on a more ad hoc basis.

Second, surveys of public servants can be costly in terms of staff time in a way that repurposing administrative data is not. Such surveys often sample a census of (that is, all) government employees. The staff time cost of completing the US federal government employee survey reaches US\$30 million annually (as cited in chapter 20). It is therefore important to design such surveys to be as efficient and effective as possible.

Third and more important, many key features of public administration production cannot be measured efficiently through other data (such as administrative data or citizen surveys). For example, understanding how public servants are managed, their motivations, and their behaviors are all internal to the official's lived experience, yet matter for public sector productivity. Public employees' motivations are difficult to observe outside of their own expressions of their motives. Thus, self-reporting through surveys becomes the primary means of measurement for many aspects of the public administration production function, and serves as a lever for improving public sector productivity.

For all these reasons, effectively designing, implementing, and making the most of surveys of public servants for government improvement is crucial.

## Surveys of Public Servants: The Global Landscape

In chapter 18, Ayesha Khurshid and Christian Schuster review current government practices in surveys of public servants. The chapter finds that governments undertake such surveys with relatively similar objectives, and thus most governments tend to measure similar concepts in their surveys. These include, on the one hand, measures of core employee attitudes correlated with productivity, such as job satisfaction and engagement, commitment to the organization, and intent to remain working for the organization. On the other hand, governments measure a relatively consistent set of management practices as antecedents of these employee attitudes, such as the quality of leadership, performance management, and training.

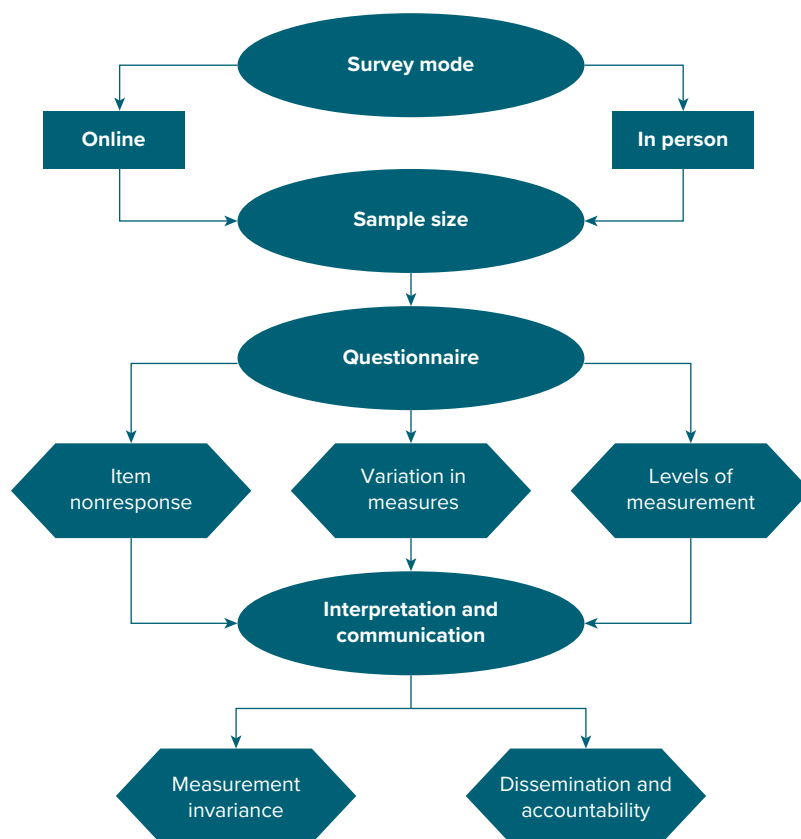
Yet chapter 18 finds that governments differ in how they design, implement, and report on surveys of public servants. For instance, the wording of survey questions differs, even when similar concepts are being measured. Approaches to sampling public servants differ, as do survey modes or approaches to dealing with nonresponse. Governments also differ widely in how they report survey results: for instance, in terms of what kind of benchmarks are reported or what levels of hierarchy inside organizations are measured and provided with results reports. Given that governments undertake surveys with similar objectives, why is there such diversity in how they design, implement, and report on surveys?

The answer, arguably, lies in part in the limited evidence available to governments that could guide choices about design, implementation, and reporting of surveys of public servants. The chapters in part 4 thus provide novel empirical evidence to enable governments and practitioners to make more evidence-based choices in response to some of these and other methodological questions in public servant surveys.

The decision tree pictured in figure 2.4 structures the choices facing governments in surveying public servants. This decision tree should *not* be read as a linear set of steps: there are interdependencies between choices. For instance, how many respondents need to be sampled depends on the expected variation in survey measures, which in turn is a function of questionnaire design.

Nonetheless, a common first choice concerns a survey mode: Are surveys conducted online, on paper, in person, by phone, or through a combination of these modes? Governments then need to determine the appropriate survey population, and an approach to sampling respondents, including determining the desired sample size given the purpose of the survey. Subsequently, questionnaires need to be designed. While measures may vary across concepts, several general concerns apply across them. For instance, how can measures be designed so that public servants are willing to answer questions (thus avoiding item nonresponse)?

**FIGURE 2.4** Decision Tree for Surveys of Public Servants



Source: Original figure for this publication.

How can measures be designed that vary sufficiently, so that comparisons between organizations or groups of public servants on these measures become meaningful? And should survey measures inquire about the individual experience of public servants themselves or ask them about their perceptions of practices in the organization as a whole?

Finally, governments need to decide how to interpret and report results. For instance, can responses from different groups—such as public servants in different countries, organizations inside a country, or demographic groups inside a country—be meaningfully compared? Or might concepts such as job engagement mean different things to different public servants (even when answering the same question), so benchmarking is not valid? Once decisions about who to benchmark are made, how can results be reported effectively to enable action? For instance, how are survey results presented to decision-makers, and who receives results? How is capacity built to enable decision-makers to take action based on survey results, and how are they incentivized to do so? The chapters in part 4 provide novel evidence on each of these key questions.

### Determining Survey Modes and Response Rates: Do Public Officials Respond Differently to Online and In-Person Surveys?

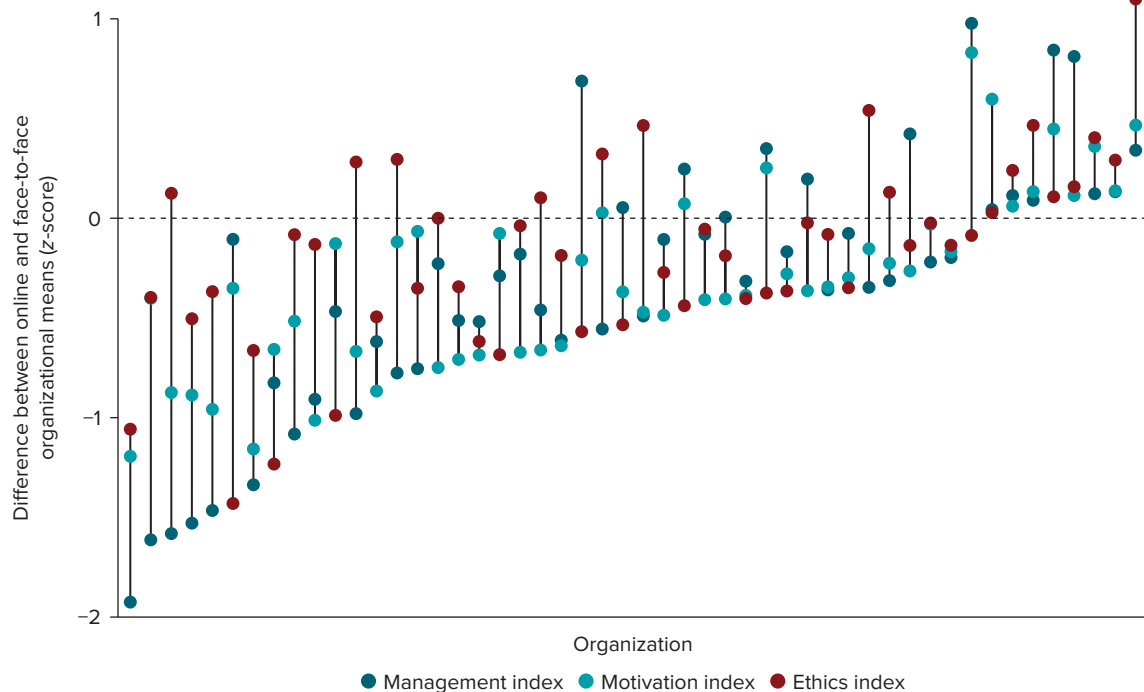
Chapter 19, by Xu Han, Camille Parker, Daniel Rogger, and Christian Schuster, assesses the first methodological choice in the decision tree: which enumeration method or survey mode to choose. This matters because different survey modes may come with different response biases to questions and different overall response rates. In OECD governments, surveys of public servants are typically conducted online, though not exclusively so. All nine government surveys reviewed in chapter 18 are implemented online, although,

to enhance accessibility (for instance, for staff who have difficulty accessing or completing an online survey), Colombia, Switzerland, the United Kingdom, and a few agencies in Australia also offer their survey in a paper format, while New Zealand offers its survey through paper and telephone upon request. The advantage of a predominantly online approach to surveying public servants across governments is clear: it reduces costs and may reduce biases, such as those induced by respondents' notions of socially desirable answers when faced with an in-person or phone interviewer.

Online surveys, however, also tend to have lower response rates than other survey modes, such as in-person surveys. For instance, the last US Federal Employee Viewpoint Survey had a response rate of 44 percent. This raises a concern that the data resulting from an online survey are not a valid representation of the population—in the case of the United States, the entire federal public administration. In some instances, these concerns about the validity of online surveys of public servants become severe. In the United Kingdom, the validity and quality of the underlying data of the Civil Service People Survey was questioned in a parliamentary inquiry, in part motivated by low response rates in some government organizations (UK Parliament 2022).

To what extent are low response rates in online surveys a concern (thus putting a premium on survey modes with higher response rates, such as in-person surveys)? To find out, chapter 19 presents evidence from a randomized control trial that compares face-to-face and online survey responses from Romanian public servants. The face-to-face surveys had consistently high response rates across Romanian government organizations, while the response rates for the online surveys varied across organizations, as is typical in other governments. The results suggest that these diverging survey modes do not substantially affect aggregate estimates at the national level. They do, however, affect the comparability of findings across organizations. Figure 2.5, reproduced from chapter 19, shows how big of a difference the mode of survey makes for indexes of survey topics. For some organizations, the impact of the survey mode is substantial.

**FIGURE 2.5 Average Difference between Survey Modes for Different Topics across Romanian Government Organizations**



Source: Original figure for this publication (see chapter 19).

Note: The figure shows, by organization in the Romanian government, the difference in the management index, motivation index, and ethics index between online and face-to-face survey respondents.



Basic organizational and demographic characteristics explain little of the variation in these effects. In other words, survey weights are not effective in addressing these effects.

Governments that offer varying survey modes should thus be careful when comparing the scores of organizations if some implement the survey primarily online while others implement it primarily based on pen and paper. Rankings of organizations in such instances do not appear to be valid. Nonetheless, chapter 19 does not find evidence to suggest that the (lower-response) online survey mode biases national-level inferences from the survey of public servants in Romania. More research is required to confirm the external validity of this finding in other countries.

In a second step, governments need to define survey populations and their sampling approach. Who the appropriate survey population is, of course, depends on the government's measurement objectives. The global review in chapter 18 shows that the survey population generally consists of central government civil servants, albeit with variations in the extent to which public sector organizations and types of employee contracts outside the (legally defined) civil service are also covered—for instance, in other branches of government or frontline services. To cite just one example, for the United Kingdom's government employee survey, all public servants from 101 agencies are eligible, excluding the Northern Ireland Civil Service, the National Health Service (NHS) (which conducts its own survey), and frontline officials (such as police officers and teachers) (UK Cabinet Office 2022).

Who, then, should be sampled within this survey population? As chapter 18 shows, approaches to sampling vary across governments, ranging from census to random, ad hoc, and stratified sampling. Australia, Canada, Ireland, New Zealand, and the United Kingdom adopt a census approach where all eligible public sector employees may participate in the survey. Switzerland and the United States use (stratified) randomized sampling approaches for most years but conduct a census every few years. Colombia's public servant survey uses a mixed approach: for larger organizations, a stratified sampling approach is used, while for smaller organizations a census is taken to protect anonymity. The Republic of Korea adopts a sampling approach for all annual surveys.

### **Determining Sample Sizes: How Many Public Officials Should Be Surveyed?**

As noted in chapter 20, by Robert Lipinski, Daniel Rogger, Christian Schuster, and Annabelle Wittels, determining the appropriate sample of a public administration survey is often a trade-off between increasing the precision of survey estimates through greater sample sizes and the high costs of surveying a larger number of civil servants. Greater precision enables both more precise benchmarking (such as between organizations) and survey result reports at lower levels of hierarchy in an organization. Less precision reduces the staff time lost responding to the survey.

How can this trade-off be resolved? Chapter 20 shows that, ultimately, survey administrators must decide on the sample size based on the type of inferences they want the survey to yield and the staff time they can claim for the survey. By employing Monte Carlo simulations on survey data from Chile, Guatemala, Romania, and the United States, chapter 20 shows that governmentwide averages can be reliably derived using sample sizes considerably smaller than those used by governments currently. On the other hand, detecting differences between demographic groups (such as gender and rank) and, in particular, ranking individual public administration organizations precisely requires larger sample sizes than are collected in many existing surveys.

These results underscore, on the one hand, the importance of not overinterpreting the substantive significance of small differences between organizations in public servant survey results (or individual ranks of organizations in results). On the other hand, the results emphasize that governments should determine sample sizes based on the type of inferences and benchmarking exercises they wish to make with the data. Existing governmental surveys of public servants do not seem to be based on such a data-driven approach to sampling. Chapter 20 addresses this gap and offers an online sampling tool to enable such sampling.

In a third step, surveys of public servants require the definition of a questionnaire. Part 4 sheds light on three cross-cutting dimensions of questionnaire design: How can measures be designed that vary sufficiently



so that comparisons between organizations or groups of public servants become meaningful? How can measures be designed so that public servants are willing to answer? And should survey measures inquire about the individual experience of public servants or instead ask them about their perceptions of practices in the organization as a whole?

### **Designing Survey Questionnaires: Which Survey Measures Vary and for Whom?**

A first prerequisite for effective question design is variation: Survey measures should provide a sufficient degree of discriminating variation across respondents to be useful—or, in other words, sufficient variation to understand differences between key comparators, such as organizations or demographic groups. Without discriminating variation across organizations, demographic groups, or countries, survey measures cannot inform governments about strengths and areas for improvement. With this in mind, chapter 21, by Robert Lipinski, Daniel Rogger, Christian Schuster, and Annabelle Wittels, assesses variation in a set of typical indicators derived from data sets of public service surveys from 10 administrations in Africa, Asia, Europe, North America, and South America.

The results show that measures related to personal characteristics such as motivation do not vary as much as those relating to management practices such as leadership. When respondents are asked to assess practices of others, such as their superior or their organization, survey responses exhibit significant discriminant variation across organizations and groups. By contrast, when respondents are asked to self-assess whether they possess desirable characteristics such as work motivation, survey responses across countries tend to be heavily skewed toward favorable answers, and variation is so compressed that meaningful differences between organizations or demographic groups are difficult to detect. Standard measures for desirable attitudes, such as motivation, may therefore need to be redesigned in surveys of public servants to better discriminate between values at the top end of indexes.

### **Designing Survey Questionnaires: To What Types of Survey Questions Do Public Servants Not Respond?**

Chapter 22, by Robert Lipinski, Daniel Rogger, and Christian Schuster, shows that surveys of public servants differ sharply in the extent to which respondents skip or refuse to respond to questions. So-called item nonresponse can affect the legitimacy and quality of public servant survey data. Survey results may be biased, for instance, if those least satisfied with their jobs are also most prone to skipping survey questions. Understanding why public servants respond to some survey questions but not others is thus important.

The chapter offers a conceptual framework and empirical evidence to further this understanding. Drawing on other survey methodology research, the chapter theorizes that public servants are less likely to respond to questions that are too complex (because they are unable to answer them) or sensitive (because they are unwilling to respond). Coding the complexity and sensitivity of public servant survey questions in Guatemala, Romania, and the United States, chapter 22 finds one indicator of complexity to be the most robust predictor of item nonresponse across countries: respondents' lack of familiarity with the information or topic examined by a survey question. By contrast, other indicators of complexity or sensitivity or machine-coded algorithms of textual complexity do not predict item nonresponse. The implication for survey design is clear: Avoid questions that require public servants to speculate about topics with which they are less familiar.

### **Designing Survey Questionnaires: Should Surveys Ask about Public Servants' Perceptions of Their Organization or Their Individual Experience?**

A third prerequisite for effective questionnaire design is valid measurement of organizational aggregates, such as which surveyed organization has the highest level of job satisfaction among its employees or

which organization has the highest quality of leadership practices of superiors. This raises the issue of whether respondents should be asked about their perceptions of organizational practice overall (so-called organizational referents) or whether questions should ask about the respondent's own experience, such as the quality of their superior's leadership practices or their own job satisfaction (so-called individual referents). In chapter 23, Kim Sass Mikkelsen and Camille Mercedes Parker examine this question using survey experiments with public servants in Guatemala and Romania. The survey experiments randomly assign public servants to respond to questions about a topic with phrasing using either an organizational referent or an individual referent.

The chapter finds that, while there are no strong conceptual grounds to prefer either organizational or individual referents—both have advantages and disadvantages—the choice matters to responses and alters response means (such as the average job satisfaction of employees in an organization). Organizational questions are particularly useful when questions are very sensitive (such as on corruption) because respondents may skew their response more strongly toward a socially desirable response option when asked about their own individual experience rather than practices in the organization. Individual questions are particularly useful when the attitudes or practices being measured are rare in the organization. In such cases, many respondents may lack the information to accurately assess the prevalence of a practice in the organization, risking that they rely instead on unrepresentative information or stories, for instance, rather than actual organizational characteristics. In short, whether survey questions should ask about the individual's own experience or the individual's perception of organizational practice depends on the characteristics of the question and the organization it seeks to measure.

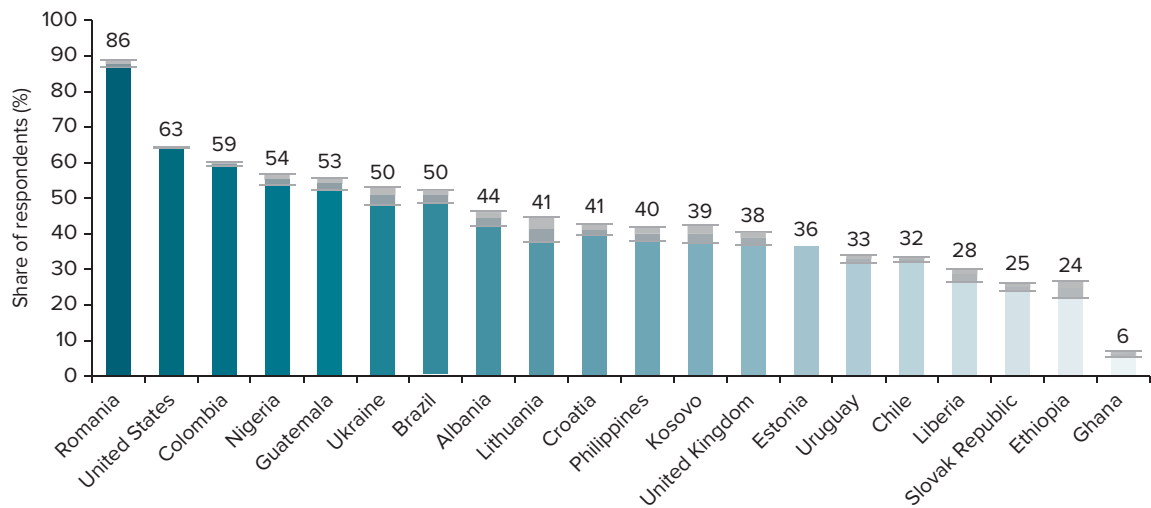
### **Interpreting Survey Findings: Can Survey Results Be Compared across Organizations and Countries?**

Chapters 24 to 26 turn to the interpretation and reporting of survey results. In chapter 24, Robert Lipinski, Jan-Hinrik Meyer-Sahling, Kim Sass Mikkelsen, and Christian Schuster focus on interpretation and in particular the question of benchmarking: Can survey results be compared across organizations and countries? This matters because survey results can rarely be understood in a void. Rather, they require benchmarks and points of reference. If, for instance, 80 percent of public servants are satisfied with their jobs, should a manager interpret this as a high or low level? Historical comparisons provide a sense of dynamics over time, but not a sense of degree. The availability of similar statistics from comparator organizations in the public sector or other countries is a potentially valuable complement to a manager's own results. However, such benchmarking requires that survey questions measure the same concept in the same way, making meaningful comparisons possible. Even when questions are phrased in the exact same way, however, the validity of comparison is not obvious. For multiple reasons, including work environment, adaptive expectations, and cultural factors, different people might understand the same question in distinct ways and adjust their answers accordingly. This might make survey results incomparable not only across countries but also across different groups of public servants within a national public administration.

To assess this concern empirically, chapter 24 investigates to what extent the same survey questions measure the same concept similarly—that is, questions are measurement invariant—using questions related to “transformational leadership” and data from seven public service surveys from Europe, Latin America, and South Asia. The chapter finds support for so-called scalar invariance: the topic (in this case, means of transformational leadership) can be compared *within* countries across organizations and demographic groups (the chapter authors test for gender and educational levels). Across countries, the chapter finds tentative evidence for scalar invariance, and stronger evidence when countries are grouped by regions and income.

The findings—although tentative and requiring further confirmatory evidence from other settings—thus underscore the utility of global benchmarking of countries in surveys of public servants.<sup>3</sup> As chapter 18 explains, the Global Survey of Public Servants offers a tool to harmonize questions across governments. In conjunction with the freely accessible Global Survey Indicators dashboard, the Global Survey of Public

**FIGURE 2.6 Share of Public Servants Satisfied with Their Pay and/or Total Benefits, Various Countries**



Source: Fukuyama et al. 2022, figure 9.

Note: Years of measurement vary by country. Colors denote the extent of job satisfaction, with darker shades signifying greater job satisfaction. The gray vertical bars denote 95% confidence intervals.

Servants thus enables governments to understand strengths and areas for development of their public administration in global comparative terms. For example, the Global Survey provides comparative data on pay satisfaction from public administrations around the world. As can be seen from figure 2.6, this varies greatly across countries.

Once decisions are made about whom to benchmark against, consideration turns to how to report and disseminate survey results—that is, how to make the most of survey results. Chapters 25 and 26 provide two complementary perspectives on this challenge.

### **Making the Most of Public Servant Survey Results: Lessons from Six Governments**

Chapter 25—by Christian Schuster, Annabelle Wittels, Nathan Borgelt, Horacio Coral, Matt Kerlogue, Conall Mac Michael, Alejandro Ramos, Nicole Steele, and David Widlake—presents a self-assessment tool that lays out low-cost actions governments can take to support evidence-based reforms based on the insights from public servant surveys. The chapter applies this tool to governments of six countries (Australia, Canada, Colombia, Ireland, the United Kingdom, and the United States) to assess the comprehensiveness of their ecosystem to turn survey results into management improvements.

The self-assessment tool focuses on three main components of an ecosystem to turn survey results into management improvements: (1) information, (2) capacity, and (3) incentives to take action. For the first component (information), public servant survey results can improve public administration by providing tailored survey results to four main types of users: the government as a whole; individual public sector organizations; individual units or departments within a public sector organization; and the public, including public sector unions. Results reporting should identify key takeaways about the strengths and weaknesses of particular organizations and enable users to explore aggregate survey results in a customized manner, such as through dashboards.

For the second component (capacity to take action), reporting of results can become more effective when it includes (automated) recommendations to users—such as managers of units or organizations—on how to best address survey findings, as well as action plans for users to develop their own actions. Where more

resources are available, tailored technical assistance—or a human resources management (HRM) consultancy, provided either by a central HR unit or an external provider—can further help managers turn survey findings into improvements.

For the third component (incentives to take action), accountability mechanisms are key. For instance, governments can introduce central oversight of actions taken in response to survey findings by government organizations and units; can construct and publicize “best place to work” in government indexes to foster external oversight; or measure employee perceptions of the extent to which government organizations take action in response to survey findings.

Applying this self-assessment framework to the six governments, chapter 25 finds that many governments could undertake a range of additional low-cost actions to enhance the benefits they derive from public servant surveys to improve public administration.

## Using Survey Findings for Public Action: The Experience of the US Federal Government

Chapter 26—by Camille Hoover, Robin Klevins, Rosemary Miller, Maria Raviele, Daniel Rogger, Robert Seidner, and Kimberly Wells—complements chapter 25 by delving into the experience of the United States, the country with the longest-standing governmentwide employee survey. The chapter emphasizes, first, the importance of considering action in response to survey results at the time the survey is being designed. In particular, questions should focus on topics that staff and senior leaders find most important to achieve their mission. Second, the chapter describes the critical architecture necessary in each public sector organization to translate survey results into management improvements. This includes, for instance, a technical expert in the organization capable of interpreting survey data; a strong relationship between that expert and a senior manager in the organization who acts as a “change champion”; and the development of a culture for initiatives for improvements informed by the survey.

The chapter also provides guidance on how to develop a culture of responsiveness to surveys of public servants. It emphasizes the importance of leaders in an organization being transparent in sharing and discussing the survey results with their workforce, codeveloping action plans with staff, and coproducing improvements in response to survey results.

Part 4 thus provides evidence-based good practice on a range of choices involved in designing, implementing, interpreting, and reporting on surveys of public servants. This evidence can inform actions by governments seeking to improve their existing regular governmentwide employee surveys, as is the case in many OECD countries. It can enable governments that have yet to introduce surveys of public servants to leapfrog to best practice from the start. Such governments are encouraged to consult the range of toolkits on the Global Survey of Public Servants toolkit site.<sup>4</sup>

## PART 5: GOVERNMENT ANALYTICS USING EXTERNAL ASSESSMENTS

Part 5 turns to select data sources available to undertake government analytics through external assessments: that is, assessments conducted on or by those outside government (rather than data on public servants or administrative data collected by government organizations themselves). Chapters 27 through 30 provide guidance on four external data sources: household survey data, citizen survey data, service delivery indicators, and anthropological methods. These data sources illustrate the possibility of government analytics through external assessments but do not cover the full range of microdata for external assessments. Enterprise surveys of businesses, for instance, can provide insights on topics such as bribery or government regulation (World Bank 2023). With that caveat in mind, the part 5 chapters provide important insights about how to do government analytics using external assessments.

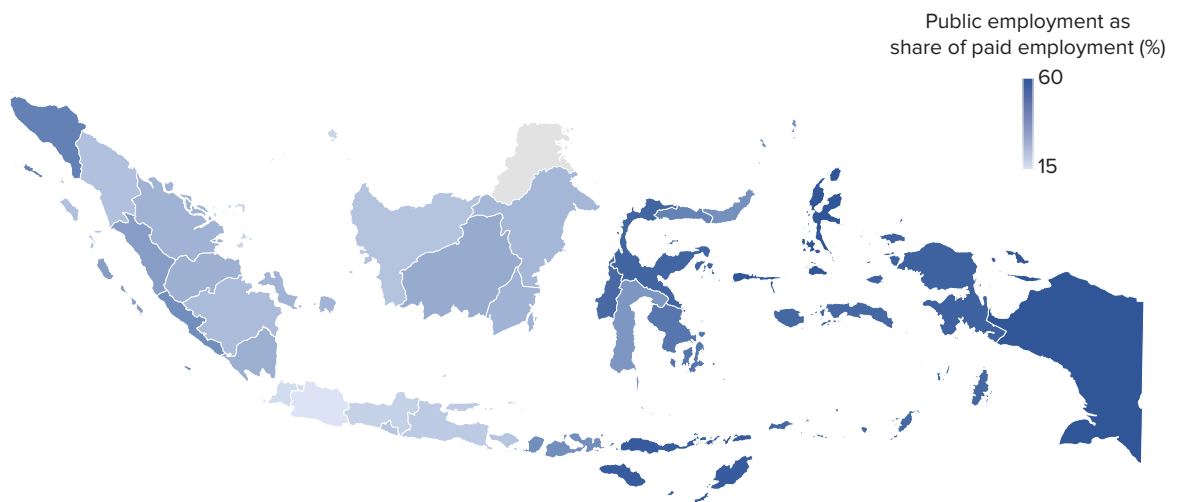
## Government Analytics Using Household Surveys

Chapter 27, by Faisal Ali Baig, Zahid Hasnain, Turkan Mukhtarova, and Daniel Rogger, describes how to use household survey data for government analytics. Such data are readily available in many countries. In particular, national statistical authorities frequently collect labor force (and related household) surveys that are broadly consistent across time and developed using globally standardized definitions and classification nomenclatures. Governments can leverage these data to gain insights into the public sector workforce that administrative data or public servant survey data do not provide. In particular, labor force survey data allow governments to explore and compare public and private sector labor markets (because labor surveys cover both populations), as well as labor markets in different regions of a country or over time. Map 2.2, reproduced from chapter 27, presents differences in labor market features for Indonesia. The role of public sector employment in the formal sector varies from 15 percent to 60 percent of paid employment, implying substantial economic vulnerability of some regions to changes in public employment practices and policies.

Chapter 27 shows how such comparisons provide a wealth of insights into the input side of the public administration production function. To cite just three examples: Labor force data analytics enable governments to understand gender pay and employment differences between the public and private sectors, and whether the public sector promotes gender equality in employment in both absolute terms and relative to the private sector. The analytics help governments assess pay competitiveness, providing answers to whether the public sector pays competitive wages compared to the private sector to attract talent while not crowding out private sector jobs. Household and labor force survey data can also shed light on the skills composition of the public sector workforce with respect to the private sector and identify in what areas the government is competing most intensively for skills with private sector actors.

In short, such data can complement payroll, HRMIS data, and public servant survey data to provide a more complete diagnosis of public pay and employment. To facilitate access to these analytics, chapter 27 also highlights the freely available Worldwide Bureaucracy Indicators (WWBI), a set of indicators based on labor force survey data from more than 200 countries compiled by the World Bank to assess public and private labor markets and their interaction across the world.

### MAP 2.2 Subnational Patterns in Public Sector Employment, Indonesia, 2018



Source: Original map for this publication, based on data from the World Bank Worldwide Bureaucracy Indicators database v 1.1, <https://www.worldbank.org/en/data/interactive/2019/05/21/worldwide-bureaucracy-indicators-dashboard>.

## Government Analytics Using Citizen Surveys: Lessons from the OECD Trust Survey

Chapter 28, by Monica Brezzi and Santiago González, examines government analytics using citizen surveys. These surveys can help governments shed light on certain outcomes in the public administration function. In particular, they can capture the outcomes of public governance as perceived and experienced by people, through nationally representative population samples. For instance, citizen surveys are used in many countries to measure satisfaction with widely used public services (such as tax administrations, schools, or hospitals).

As chapter 28 shows, they can also be used to understand broader government outcomes. In particular, the chapter illustrates the potential of such surveys for government analytics using the example of the OECD's Survey on the Drivers of Trust in Public Institutions (OECD Trust Survey). The survey measures trust of citizens in government and captures their expectations of and experiences with public institutions around key drivers of trust. Measures of trust and its drivers include the competence of public institutions—including access to public services and their quality and reliability—as well as the perceived values of public institutions, notably in terms of integrity, openness to involving citizens, and fairness in the treatment of citizens.

Chapter 28 showcases how governments have used evidence resulting from the survey to develop concrete actions to strengthen institutional trust. The chapter provides guidance for other governments wishing to apply this international benchmark on measuring trust in public institutions, following the OECD Guidelines on Measuring Trust.

## Government Analytics Using Measures of Service Delivery

In chapter 29, Kathryn Andrews, Galileu Kim, Halsey Rogers, Jigyasa Sharma, and Sergio Venegas Marin go beyond this book's core focus on public administration to the frontline of service delivery and introduce measures of service delivery (MSDs). Currently, MSDs are focused on education and health. Mirroring this *Handbook's* approach to conceptualize government analytics along the public administration production function from inputs to outcomes, MSDs provide objective measurements not only of service quality (such as absenteeism of medical doctors, and test results of students in school) but the entire process involved in delivering frontline public services (including input and process measures such as availability of medicine, and management practices in schools).

MSDs offer a granular view of the service delivery system, providing actionable insights on different parts of the delivery chain, from the physical infrastructure to the knowledge of frontline providers. These can be usefully viewed as outputs and “outcomes” of the administrative environment embedded in government agencies under which these frontline providers fall. Measurement of these different factors of production allows practitioners to map out, conceptually, how each part of the production chain is faring, and where improvements can be made, at the individual provider level as well as part of the wider production function for government outlined in figure 2.1. MSDs also provide action-oriented visualizations of these indicators, enabling practitioners to design their service delivery policies in an intuitive and evidence-based approach.

Chapter 29 provides a road map to interested practitioners to produce MSDs, from design, implementation, and analysis to dissemination. The chapter emphasizes that developing service delivery indicators requires considering and defining relevant dimensions of quality in a public service, along with relevant inputs, with a premium on defining indicators according to policy objectives and resource constraints. Akin to the discussions in chapters 25 and 26, chapter 29 also underscores the importance of linking MSDs directly to stakeholders who have the ability to enact change in the delivery system. Many of the steps involved in the analytics of core public administration are thus mirrored in the analytics of service delivery.



## Government Analytics Using Anthropological Methods

Finally, chapter 30, by Colin Hoag, Josiah Heyman, Kristin Asdal, Hilde Reinertsen, and Matthew Hull, returns to chapter 4's call for embedding qualitative studies in government analytics. It considers how government analytics can be undertaken through an anthropological approach, a powerful means of collecting qualitative data. Anthropologists are most commonly associated with immersive, ethnographic methods such as participatory observation. Chapter 30 applies that lens to studying public administration. As the chapter authors emphasize, "Anthropologists are motivated by an abiding concern with empirical rigor—a refusal to ignore any sort of data or to content oneself with a single view of such a multifarious thing as bureaucracy." Doing so risks overlooking factors that shape organizations.

Anthropological methods suggest that data collection should approach government analytics by engaging with the staff who are involved at every level of the organization, from senior officers to low-level staff and contractors, and across different demographic groups; studying everyday documents; and watching how officials interact. By observing every part of what public officials do at work in a holistic way, from their interactions in corridors and meetings to the protocols they observe in their relationships, the analyst undertakes the most holistic data collection strategy feasible.

Such an approach requires analysts to develop relationships with a variety of types of people in an organization and have open-ended conversations about their work and unrelated issues to understand their values and perspectives. It also requires analysts to engage in participant observation to capture activities that may be so routine they go unnoticed by public officials and are not self-reported in surveys. Moreover, it requires analysts to collect the widest practical range and amount of qualitative and quantitative data, even if such data cannot be easily standardized. Finally, it requires analysts to study not only data but also the interactions and microscopic decisions that affect the gap between stated policy goals and the actual work being carried out by public officials—for instance, by studying what public officials say and do, including the rationales they draw on for their decisions.

Chapter 30 thus emphasizes that government analytics can incorporate methods that provide insights into aspects of the public administration function that quantitative microdata cannot. It also brings the *Handbook* back to the first cross-cutting chapter (chapter 4), which emphasizes the importance of a holistic and "balanced data suite." Part of this "suite" is ensuring that problem analysis in public administration is holistic: that important parts of a problem are not neglected due to the absence of quantitative data and measurement. This, in turn, puts a premium on utilizing qualitative and anthropological methods to complement insights gleaned from microdata.

Holistic analytics benefit not only from triangulating different quantitative and qualitative methods, but also from triangulating and integrating the analytics of different components of the public administration production function. The chapter concludes with a challenge: How can the different approaches and data sources in the government analytics toolbox be integrated effectively to diagnose major challenges in public administration holistically?

## CONCLUSION: TOWARD A HOLISTIC ANALYTICS OF CORE CHALLENGES IN THE MACHINERY OF GOVERNMENT

How can practitioners take the distinct data sources detailed in the *Handbook* to the frontier? Frontier government analytics would integrate the analytics of the data sources described across the *Handbook* into standard government practice. It would generate them at a scale sufficient to inform the decision-making of individual managers. And it would make them easily accessible to those managers across government organizations and departments. For instance, dashboards integrating data sources and updating in real time would provide managers with comparisons for their staffing issues, process quality, the perceived quality of



management practices, and so on. They could keep tabs on outputs and outcomes, from task completion and case productivity to external assessments from citizens. Comparative data would allow them to benchmark themselves against other government organizations, or where appropriate, other countries. Managers would be capable of understanding the limitations and strengths of different analytics. The result would be a transformational change toward leveraging data to strengthen public administration.

Where should the journey toward this frontier begin? As a first step, the government analytics of the individual data sources explored in detail in various chapters can provide important insights into different components of the public administration production function. Once governments have analytical tools for several components in place (such as payroll diagnostics, and surveys of public servants), the possibilities for government analytics further expand. The integration of multiple government analytics data sources enables governments to diagnose and address major government challenges holistically (chapter 9). In some cases, such as customs (chapter 14), such integration is a vital part of measuring attainment of the organization's mission. In others, the insights that can be gleaned go beyond those available from individual data sources and thus enable a holistic perspective on public administration.

To illustrate, consider how the integration of government analytics data sources described in this *Handbook* can shed light on key challenges in public administration: corruption and personnel management.

Corruption is a multidimensional phenomenon, affecting public administration across its production function. As chapter 8 discusses, corruption can first be detected with input data. For example, payroll data can be drawn on to detect ghost workers. HRMIS data can be drawn on to detect instances of nepotism in recruitment, such as when family members with similar last names are hired. Procurement data can be drawn on to detect procurement fraud and collusion risks (such as when organizations grant contracts without competitive bidding). Expenditure data can be drawn on to detect off-budget spending at risk of embezzlement.

Second, corruption can be detected in the processes and practices that convert inputs into outputs and generate norms and behaviors in public administration. For instance, surveys of public servants can measure unethical leadership by superiors (such as pressures on subordinates to collude in corruption), as well as the ethical norms and integrity of public servants themselves (such as their perceptions of corruption of colleagues).

Third, corruption can be detected in output and outcome data (such as in tax audit case data, or through surveys of citizens querying them about bribery requests from public officials). Understanding at a granular level where corruption occurs—in a specific public administration production function, or in a particular department or organization—enables governments to identify comprehensive but tailored and evidence-based solutions to curb corruption.

Consider, as a second example, civil service management—and how the integration of analytics across the public administration production function can aid better personnel management in government. As chapters 9 and 10 emphasize, data analytics of payroll and HRMIS systems can diagnose personnel as an input into the production function. To cite two examples: pay equity can be assessed across groups (such as by gender) and institutions; and retention challenges can be pinpointed by comparing turnover rates for public servants at different ranks or in different regions of the country. Such data can be complemented by labor force and household survey data (chapter 27)—for instance, to understand whether public sector pay is competitive relative to the private sector, or whether the workforce allocation across the territory is appropriate given the respective number of service users. Other input data, such as on budgets, can help understand whether productivity problems might arise from the lack of complementary inputs to personnel in the production function.

Data on processes and practices can then shed light on whether government is effectively converting personnel inputs into outputs and outcomes. Surveys of public servants, as explored in part 4, can measure the experience of public servants with management practices, as well as the culture, attitudes, and behaviors in public administration that mediate the conversion of inputs into outputs. For instance, is leadership by line managers effective? Are public servants motivated to work hard? Anthropological and qualitative

methods, as discussed in chapter 30, can enrich and contextualize survey responses through participant observation and detailed case study work. As chapter 13 shows, government analytics can also assess processes converting inputs into outputs through administrative data—for instance, to assess whether institutions follow due procedure when evaluating the performance of public servants.

Last, data on outputs and outcomes can help diagnose and improve personnel management. For instance, as chapter 17 explains, data on task completion can help understand where in government public servants are effectively completing tasks and where they are not. For public sector organizations that complete cases, administrative case data, as explored in chapters 14 and 15, can help understand differences in the productivity of public servants more broadly, while surveys of citizens can help understand differential citizen satisfaction and trust with public services across institutions and regions in a country.

### Strengthening Public Sector Management through Government Analytics

At the center of government are the public officials who navigate the strategic and daily decisions that determine public policy and the effectiveness of public administration. The information public managers have, and the extent to which they use it for running public organizations, drives the efficacy of government. An effective system of government analytics empowers officials to manage government based on evidence of the administration's current realities.

Empowering public officials with government analytics can transform the basis for personnel management—and public administration more generally. It enables government decision-makers to complement their tacit and practical knowledge about public administration with evidence-based and data-informed insights to improve how the machinery of government is run.

Almost all officials manage, organize, or process in ways that could usefully be analyzed as an input to strengthening government functioning. Thus, communicating analytical insights to government officials at the time that they are making a decision, as discussed in chapters 25 and 26, is a critical link in making analytics effective. One approach is to support decision-making through a human resources management dashboard that brings together the distinct data sources a government has available to visualize and benchmark the strengths and areas for improvement in personnel management to decision-makers in each government institution. By implementing any of the approaches in this *Handbook*, government or its stakeholders are building a platform for managers and decision-makers to do more with the public sector's finite resources.

Even with 30 chapters, this book does not cover all potential data sources and approaches to government analytics. There will always be new analytical opportunities on the horizon. Chapter 3 discusses how government can continuously prepare for a transition to new analytical approaches. Above all, building the culture that binds public administrators and public administration together requires a commitment from senior management and individual officials to use evidence about their own administration in their decision-making. We invite all public servants and related stakeholders to capitalize on the insights summarized in this chapter and this *Handbook* and push forward the quality of management of their administration. The quality of government around the world will be shaped by how its decision-makers leverage data to strengthen their administration.

## ANNEX 2A MAPPING GOVERNMENT ANALYTICS DATA SOURCES TO CHAPTERS IN THE HANDBOOK

**TABLE 2A.1 Mapping Government Analytics Data Sources to Chapters 10–30**

Data source	Chap.	Examples of uses of analytics	Examples of indicators
<i>Part 3. Administrative data</i>			
Payroll and HRMIS	10	<ul style="list-style-type: none"> <li>Examine fiscal planning and sustainability of the wage bill.</li> <li>Allocate the workforce across government departments and territories.</li> <li>Set pay.</li> </ul>	<ul style="list-style-type: none"> <li>Wage bill by sector and years.</li> <li>Distribution of civil servants by rank.</li> <li>Turnover of civil servants.</li> <li>Pay inequity between ministries.</li> <li>Retirement projections.</li> </ul>
Expenditure	11	<ul style="list-style-type: none"> <li>Assess efficiency, equity, and effectiveness of government spending.</li> <li>Assess whether government spending corresponds to budget priorities.</li> <li>Identify large transactions with high fiduciary risk.</li> </ul>	<ul style="list-style-type: none"> <li>Budget execution rates.</li> <li>Share of government expenditures covered by the FMIS.</li> <li>Share of total expenditures by transaction value.</li> <li>Overdue accounts payable (payment arrears).</li> </ul>
Procurement	12	<ul style="list-style-type: none"> <li>Monitor procurement markets and trends.</li> <li>Improve procurement and contracting (for instance, by identifying goods organizations overpay for, or organizations with high corruption risks in procurement).</li> <li>Assess effects of distinct procurement strategies or reforms.</li> </ul>	<ul style="list-style-type: none"> <li>Time needed for contracting.</li> <li>Number of bidders.</li> <li>Share of contracts with single bidder.</li> <li>Final price paid for a good or service.</li> <li>Share of contracts with time overruns.</li> <li>Share of bidders that are small and medium enterprises.</li> </ul>
Administrative processes	13	<ul style="list-style-type: none"> <li>Assess the speed and quality of administrative back-office processes and process implementation (for instance, for project planning, budget monitoring, or performance appraisals).</li> </ul>	<ul style="list-style-type: none"> <li>Adherence of administrator's process work to accepted government procedure.</li> <li>Timeliness of administrator's process work with respect to deadlines.</li> </ul>
Customs	14	<ul style="list-style-type: none"> <li>Assess customs revenue collection.</li> <li>Assess trade facilitation (flow of goods) through customs across borders.</li> <li>Assess whether customs safeguards safety of goods and protects people (for instance, prevents dangerous goods from crossing).</li> </ul>	<ul style="list-style-type: none"> <li>Time delays in customs clearances.</li> <li>Cost of customs process.</li> <li>Total customs revenue collected.</li> <li>Number of goods in infraction seized in customs.</li> </ul>
Case data	15	<ul style="list-style-type: none"> <li>Assess productivity of organizations and individuals processing cases (such as tax claims).</li> <li>Assess the quality of case processing.</li> <li>Identify "best performing" offices and transfer best practices to other offices.</li> </ul>	<ul style="list-style-type: none"> <li>Total number of cases completed.</li> <li>Average time to complete one case.</li> <li>Error rate in completing cases.</li> <li>Timeliness of case completion.</li> <li>Share of cases with complaints.</li> </ul>

(continues on next page)

**TABLE 2A.1 Mapping Government Analytics Data Sources to Chapters 10–30**  
(continued)

Data source	Chap.	Examples of uses of analytics	Examples of indicators
Text-as-data (machine learning)	16	<ul style="list-style-type: none"> <li>Analyze “big data” and convert unstructured text from government records into data for analytics.</li> <li>Examples of big data include payroll disbursements to civil servants or tax filings by citizens and firms.</li> <li>Examples of text data include court rulings, procurement tenders, or policy documents.</li> </ul>	<ul style="list-style-type: none"> <li>Risk score for a procurement tender or payroll disbursement based on a predictive algorithm.</li> <li>Bias score in court ruling, based on textual analysis of wording in the document (sexist or racial-profiling terms).</li> </ul>
Task and project completion data	17	<ul style="list-style-type: none"> <li>Examine the frequency of completion of tasks on schedule (such as delivery of training program, preparation of budget).</li> <li>Understand quality of administrative task completion.</li> <li>Understand drivers of differences in task completion across government units.</li> </ul>	<ul style="list-style-type: none"> <li>Number of tasks completed.</li> <li>Share of tasks completed on time.</li> <li>Share of unfinished tasks.</li> <li>Share of tasks completed as planned.</li> </ul>
<i>Part 4. Surveys of public servants</i>			
Civil service survey data	18–26	<ul style="list-style-type: none"> <li>Assess the quality of management practices (such as performance management, recruitment).</li> <li>Assess norms, attitudes, and cultures in public administration (such as work motivation).</li> </ul>	<ul style="list-style-type: none"> <li>Share of public servants hired through merit examinations.</li> <li>Share public servants wishing to leave the public sector in the next year.</li> <li>Share of public servants favorably evaluating the leadership practices of their superior.</li> </ul>
<i>Part 5. External assessments</i>			
Household survey data	27	<ul style="list-style-type: none"> <li>Assess wage differences between the public and the private sector.</li> <li>Identify regions with lower public sector wage premiums.</li> <li>Assess gender representation in the public sector.</li> </ul>	<ul style="list-style-type: none"> <li>Average public sector wage premium.</li> <li>Share of women in public employment.</li> <li>Share of public employment in total employment.</li> </ul>
Citizen survey data	28	<ul style="list-style-type: none"> <li>Assess satisfaction of citizens with public services.</li> <li>Assess trust of citizens in government institutions.</li> <li>Assess interactions of citizens with public administration (such as bribery requests).</li> </ul>	<ul style="list-style-type: none"> <li>Share of citizens satisfied with health services.</li> <li>Share of citizens trusting the civil service.</li> <li>Share of citizens asked for a bribe by a government official in the last year.</li> </ul>
Service delivery indicators	29	<ul style="list-style-type: none"> <li>Assess the quality of education and health service delivery.</li> <li>Assess facility-level characteristics (such as teacher presence).</li> </ul>	<ul style="list-style-type: none"> <li>Share of days during which teachers are not present in school.</li> <li>Availability of resources in hospital to treat patients.</li> </ul>
Anthropological analytics	30	<ul style="list-style-type: none"> <li>Observe everyday practices in public administration, to capture routine but unnoticed parts of administration.</li> <li>Observe social engagement to understand formal and informal rules, relationships, and interactions between public servants.</li> </ul>	<ul style="list-style-type: none"> <li>Holistic participant observation of everyday life inside particular public administrations.</li> <li>Understand how public servants interpret broader policy goals.</li> </ul>

Source: Original table for this publication.

Note: FMIS = financial management information system; HRMIS = human resources management information system.

## NOTES

1. In contrast to the relatively coherent consensus of functions of private sector production (Mas-Colell, Whinston, and Green 1995), no consensus has formed around an integrated model of a production function for public administration. This is due in part to the limited use of microlevel data on the workings of different components of public administration—the very gap that motivated this *Handbook*.
2. Whether inputs effectively convert into outputs is also moderated by exogenous factors, such as the political environment. This *Handbook* does not discuss the analysis of microdata to assess these exogenous factors. Instead, readers are encouraged to consult the many existing excellent resources to understand the exogenous and political environment of public administrations (see, for example, Moore 1995).
3. Read in conjunction with other research cited in chapter 24, this conclusion holds particularly for questions that are more factual and less culturally specific. For instance, questions on specific management practices (such as around the presence of certain performance evaluation practices) can be more plausibly benchmarked across countries without measurement invariance concerns than attitudinal questions (such as on how engaged employees are).
4. The Global Survey of Public Servants was cofounded by the two editors of this *Handbook*, along with a range of practitioners and academic colleagues. The toolkits on the website build upon and incorporate much of the evidence reviewed in this *Handbook*. See <https://www.globalsurveyofpublicservants.org>.

## REFERENCES

- Bakshi, H., A. Bravo-Biosca, and J. Mateos-Garcia. 2014. “The Analytical Firm: Estimating the Effect of Data and Online Analytics on Firm Performance.” Nesta Working Paper 14/05, Nesta, London.
- Fenzia, A. 2022. “Managers and Productivity in the Public Sector.” *Econometrica* 90 (3): 1063–84.
- Fukuyama, F., D. Rogger, Z. Husnain, K. Bersch, D. Mistree, C. Schuster, K. Sass Mikkelsen, K. Kay, and J.-H. Meyer-Sahling. 2022. Global Survey of Public Servants. <https://www.globalsurveyofpublicservants.org/>.
- Mas-Colell, A., M. D. Whinston, and J. R. Green. 1995. *Microeconomic Theory*. Oxford, UK: Oxford University Press.
- Meyer-Sahling, J., D. Mistree, K. Mikkelsen, K. Bersch, F. Fukayama, K. Kay, C. Schuster, Z. Hasnain, and D. Rogger. 2021. *The Global Survey of Public Servants: Approach and Conceptual Framework*. [https://www.globalsurveyofpublicservants.org/\\_files/ugd/acee03\\_91fc43ed92774eb88704717d6c7c80a2.pdf](https://www.globalsurveyofpublicservants.org/_files/ugd/acee03_91fc43ed92774eb88704717d6c7c80a2.pdf).
- Moore, M. 1995. *Creating Public Value: Strategic Management in Government*. Cambridge, MA: Harvard University Press.
- UK Cabinet Office. 2022. “UK Civil Service People Survey: Technical Summary.” <https://www.gov.uk/government/publications/civil-service-people-survey-2022-results/civil-service-people-survey-2022-technical-guide>.
- UK Parliament. 2022. “MPs to Probe Quality of Civil Service People Survey.” <https://committees.parliament.uk/committee/327/public-administration-and-constitutional-affairs-committee/news/173010/mps-to-probe-quality-of-civil-service-people-survey/>.
- UN (United Nations). 2014. *A World That Counts: Mobilising the Data Revolution for Sustainable Development*. Report prepared at the request of the United Nations Secretary-General by the Independent Expert Advisory Group on a Data Revolution for Sustainable Development. <https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>.
- World Bank. 2023. Enterprise Surveys. <https://www.enterprisesurveys.org/en/enterprisesurveys>.
- World Bank Group. 2019. “Innovating Bureaucracy for a More Capable Government.” <https://openknowledge.worldbank.org/handle/10986/31284?show=full>.





## CHAPTER 3

# Government Analytics of the Future

*Daniel Rogger and Christian Schuster*

### SUMMARY

The investments governments make in measurement today will determine what they know tomorrow. Building an analytics system in government has long-term benefits for our ability to manage scarce public resources and detect unforeseen risks. This chapter provides guidance on what public organizations can do today to become more analytical tomorrow. Government institutions themselves require reshaping: by enhancing structures for planning; by equipping public sector managers with a greater ability to consume and interpret analytics; and by developing new architecture for analytical units. Assisting each public sector organization to develop its own analytics agenda induces cultural change and targets the analytics to the requirements of its specific mission. Rewarding experimentation with novel data sources improves government's capacity to innovate more generally. Each of these changes helps chart a course to the government analytics of the future.

### ANALYTICS IN PRACTICE

The guidance that follows aims to facilitate the transition process to build an environment for analytical insights across government:

1. Continuously plan to capitalize on the opportunities afforded by innovations in measurement and analysis of government functioning.
2. Develop units of government analytics at the center of government and within each major organizational unit, and embed them in a community of practice. Centralized units enable economies of scale in both the implementation of analytics and the breadth of comparable data created, as well as network economies from users investing in a common data architecture. Units within organizations can complement central analytics by helping interpret analytics for their organization, and adapting analytics tools to particular organizational needs.

3. Build a public sector management cadre able to undertake and interact with frontier measurement and government analytics. Technological advances in measurement and analysis reinforce the importance of capable public sector managers. A cadre of public managers literate in government analytics is aware of the boundaries and assumptions of government measurement and analysis and adept in using analytical results to complement a broader understanding of public service.
4. Pursue a centralized analytical agenda that harmonizes common variables and conducts joint analysis of them, with specific agencies supporting this “public good.” The lack of objective benchmarks in many areas of government work puts a premium on harmonization and benchmarking through common variables across organizations. Similarly, governments should invest in measures that allow international comparisons.
5. Incentivize experimentation and an innovation culture in analytics through institutions that take responsibility for failures. Cultural shifts that reward smart experimentation irrespective of the outcome often come from the explicit support of senior leadership and political actors who endorse the process of innovation—and corresponding success and failure. To reinforce that cultural shift, actors across government—from senior managers through unit supervisors to individual employees—should define analytics agendas for their areas of responsibility.

## A GOVERNMENT COMMITMENT TO CONTINUOUSLY SEEK INNOVATION

The choices governments make today on what aspects of their machinery to measure and how to do so will determine what governments know tomorrow. Reviewing the analytical status quo, and planning for its future, should be an integral part of organizational and governmentwide strategies. This chapter provides guidance on how to build a government analytics of the future based on lessons in the chapters of *The Government Analytics Handbook*. Its starting point is recognition that measurement challenges are a defining feature of the public sector.

Outcomes of government intervention in many areas of the economy and society are often hard to observe. Inside government, measurement challenges pervade public management, with managers engaging their staff in tasks that cannot be fully defined in a manual or contract and that can change rapidly in response to a societal or political shock. Management in government is thus anchored in ambiguity and uncertainty, rather than measurement and measurability. This is the environment that public sector managers must make their decisions in every day.

This has always been true. Early governments grappled with measuring the scale of the economy and its taxable component. As governments have scaled up their activities to take on an increasing role in society, they have had to measure an increasingly broad range of activities. As the complexity of society grows, so does the complexity of the organizations government must build to manage its interactions with society, and the corresponding measurement tasks. Conversely, given the centrality of the government budget process, public sector managers have had to collapse much of their measurement and activity back to cash terms that can be put into a centralized budget. Thus, public officials have always faced the tension between the incompleteness of what they know and the practical requirement to make policy decisions.

Take the case of regulation. The performance of public sector regulators will be judged by regulators’ capacity to understand and make effective decisions about the sectors that they regulate. As society’s economic environment becomes more complex, it naturally yields more complex issues for regulators to comprehend. In response, governments may hire more specialized public sector professionals to undertake required analysis of the complexity—which in turn increases the complexity of the public sector itself. Government must then determine the performance of those professionals, how they should be managed, and what might increase their capacity to undertake their job now and in the future.

Thus, in the future, government will struggle with measurement issues, just as its predecessors have. But there is a qualitative difference: in terms of both its understanding of the world it must govern, and in comprehending its own structures, capabilities, and challenges, the future of government will be far more complex as the world grows more complex.

Fortunately, however, future efforts will also benefit from technical advances in three aspects that facilitate the measurement and management of government, as discussed in chapters 1 and 2. The first is the digitization of many government services, activities, and records. The second is improvements in the analytical technology for understanding government functioning. The third is the increasing redesign of government as an analytical organization. An organization that is not designed for the collection and use of analysis about its functioning will simply not be able to use evidence, however rich. Such a redesign is multifaceted, starting with increased recognition of the need for more and better analytics, continuing through building the mindset and skills of public officials to support an analytical approach, and gaining momentum by setting up institutions that can undertake analytics and embed it in public administration.

In other words, the future of government is characterized by tension. On the one hand, governments can capitalize on the opportunities that innovations in digitization and analytics afford. On the other hand, they face increasing complexity in both society and in the organization of government. Managing the interplay between these two forces will be a central challenge. Government must build its capacity to engage with greater complexity in the world and within its own architecture.

How can this be done? This chapter provides answers. It begins by describing the institutional architecture that has been shown in the cases covered by the *Handbook* to strengthen government analytics. It continues by outlining what an effective agenda might look like to capitalize on the analytics that are available today and what may be available in the future. It concludes with a discussion of how a government's analytical architecture and analytical agenda might prepare for novel data sources that are not yet part of the administration's strategy but could be a useful addition in the future. These transformations require an informed conversation throughout the public service that drives the requisite cultural change. This *Handbook* aims to improve that conversation.

## **BUILD A GOVERNMENT OF THE FUTURE**

### **A Vision of the Future**

Like any forward-planning activity in the public sector, government should continuously plan to capitalize on the opportunities afforded by innovations in measurement and analysis of government functioning. Given the speed at which analytics is evolving, this should be a constant task. Each new innovation in the measurement of government functioning is a new opportunity for improved public sector performance.

The future of government analytics thus begins with an approach by government to continuously support innovation. This should become a strategic goal of politicians in their political agendas; of senior management of the public service in their guidelines for public service action, formulation of appraisals, and circulars to staff; and of middle management in their prioritization of analytical activities and use of corresponding results in their decision-making. Implementing these commitments in planning documents and work plans presents stakeholders with a credible signal of cultural change. These efforts are catalyzed by the formation of coalitions of political and technical officials interested in developing or innovating the analytics agenda. Chapter 26 describes how such coalitions around the Federal Employee Viewpoint Survey have substantively improved the US federal government.

Planning that includes a review of analytical opportunities should capitalize on the best evidence available in public decision-making. For example, in workforce planning, basic analytics would regularly monitor the likely shortfalls in staffing as current employees leave or retire. A higher level of analytics would aim to

predict what new roles might be necessary and which might become redundant. An even higher level would assess trends in the productivity of distinct roles, enabling adjustments to be made for shifting burdens of work. Attaining each of these levels entails strategic choices in bringing empirical evidence to the management of the public service. It also requires resources to be allocated for analytical purposes, and necessitates technical staff to work closely with senior managers to articulate the implications for personnel management. Chapter 10, for instance, zeroes in on how investments in the use of personnel management data have allowed future wage costs to be predicted. The chapter describes a ladder of quality of data, with each higher layer enabling an increase in the depth of analytical insights. The analytics provided a platform for the corresponding governments to head off fiscal shortfalls and avert a major wage bill crisis. But it was the strategic choices key decision-makers made to request the analysis and proactively respond to the results that were the key to success.

Public service cultures often guard against rapid innovation and technological change. To enable cultural change, a multitude of approaches can be taken. For instance, to signal high-level support, senior management can convey a vision of a government managed based on evidence. Senior managers and middle managers can celebrate the implementation of analytical approaches. Publicizing and explaining how government functioning in a specific agency has improved can help shift service norms toward acceptance. Hiring employees trained in data analytics and upskilling existing employees in data analytics can increase the interest in adopting innovations in analytics, and reduce the cost of making those changes. Evidence of how quickly public service culture can accept—and come to expect—novel measurement of administrative functioning can be seen in the rapid adoption of surveys of public servants in many countries in recent years, documented in chapter 18.

### Creating or Expanding Management Capabilities

As the task of integrating precision analytics with less measurable aspects of government work becomes more sophisticated, the need will grow for decision-makers capable of interpreting and integrating analytical insights with tacit managerial knowledge. For example, in the case of machine learning (ML), chapter 16 notes that “continuous collaboration between the ML implementation team and policy colleagues who will use its insights ensures that applications are adapted to and stay relevant to public administration’s needs.”

Ethical considerations are also paramount. For instance, chapter 16 emphasizes the important role public managers must play in assessing the ethics of machine-learning approaches in specific cases. Balancing the need for innovation to collect and analyze more and better data and safeguarding the public good will always be a fundamental aspect of public managers’ application and oversight of analytics. This is particularly true when data concern government itself because managers and organizations are the ultimate safeguards of their employees’ rights. Yet, as chapter 6 notes, “there is a dearth of discussion and practical guides on the ethics of data collection by government on its own employees.” The chapter provides a framework for public sector managers to judge the ethical use of government analytics in particular cases.

Another important foundation is to build what chapter 4 calls a balanced data suite to inform decision-making. As the chapter warns, “an overreliance on quantitative data comes with its own risks, of which public sector managers should be keenly aware.” While embracing a role for qualitative data, especially for those aspects that require in-depth, context-specific knowledge, analytics should focus quantitative measures of success on those aspects that are close to the problem. Analytics also needs to protect space for judgment, discretion, and deliberation in those (many) decision-making domains that inherently cannot be quantified. One way to attain a balanced suite is through the use of external assessments, as discussed in part 5, such as anthropological methods (chapter 30). To attain balance in the data suite, public managers need to identify and manage the organizational capacity and power relations that shape data management.

Managers of the public service also need to be able to have an informed discussion about when measurement is of the right nature and accuracy to make a particular claim. For example, chapter 20 shows that public servant surveys frequently do not have a sufficiently large sample to make valid comparisons across organizations about employee attitudes. An informed public official could therefore disregard such comparisons

when there is not a statistical basis to make them. The more profound the understanding of public sector decision-makers as to how measurement should be undertaken and how related analysis should optimally be mapped into decisions, the more useful government analytics will be.

All this implies that along with a commitment to analytical work, building a government of the future requires building a public sector management cadre capable of directing and interacting with frontier measurement and government analytics. This cadre should be aware of the boundaries and assumptions of that measurement and analysis and be capable of using analytical results in the context of their broader tacit understanding of the public service. Such managers should also be continuously aware of what the frontier of good practice looks like in undertaking analytical work. As chapter 5 shows, the range of freely available resources to support achieving this awareness is expanding rapidly.

Managers in individual organizations need to link to a community of practice, where they can combine learning from their own organization—and from their specific set of tasks—with learning from others. Embedding public managers in a community of practice for government analytics across the relevant administration, or across government, bolsters opportunities for learning and motivating officials, rather than leaving them as independent analysts who could be subsumed within the wider institutional environment. The network effects that arise from such a community underlie the rationale for central offices of government analytics. To encourage network effects, for instance, the US federal government holds workshops to build communities and connect with staff (see chapters 9 and 26).

## **Analytics Architectures**

Centralized units of analytics enable economies of scale in both the implementation of analytics and the breadth of comparable data created, as well as facilitating network economies from users investing in a common data architecture. For example, by mainstreaming public servant surveys into an integrated data system, a single entity can brand, market, and implement the survey (chapter 25); the statistical rigor of question design and analysis can be improved (chapters 19 through 23); and all agencies and units can compare their results to similar entities across the service (chapter 24). As more agencies use the survey for personnel management, the cultural norms around acceptability of the usefulness of the survey results shift and favor adoption (chapter 26).

Such benefits might be realized by mainstreaming government analytics into Integrated National Data Systems, typically managed by national statistical agencies. Locating analytics teams in statistical authorities may improve the statistical rigor and country ownership of the corresponding analysis. Such agencies provide a solid foundation for achieving scale and sustainability in the collection of data on public service. They also offer a platform for integrating data on government functioning with broader data on the targets of government action, such as the Sustainable Development Goals.

However, locating analytics teams in national statistical agencies outside of management agencies risks isolating analytics teams from decision-makers. In particular, these teams may not be responsive to the requirements of specific managers. To address that issue, the UK Cabinet Office and the US Office of Personnel Management have created centralized units of government analytics and located them in central management authorities rather than statistical agencies. Centralized delivery approaches have typically housed analytics teams within the heart of government, either in the presidency/prime minister's cabinet office or in a ministry of finance or public administration. The evidence leans toward developing units of government analytics at the center of government and within each major organizational unit, though whether this holds in a given government depends on its institutional context.

There may be ways to share analytical responsibilities across national statistical authorities and implementing agencies, but at this nascent stage in government analytics, few examples of such relationships exist. One example is Colombia's National Statistical Office (DANE), which conducts the country's governmentwide employee survey (chapter 18). Statistical agencies can also use existing data, such as household surveys, to provide insights into the functioning of public administration (see chapters 27 and 28). Chapter 29 provides examples of how some sectors and associated line ministries can use service

delivery assessments as diagnostic tools, particularly when combined with other forms of data on public administration, as discussed in chapter 8.

Having distinct analytics teams spread across a public service carries its own risk: fragmentation in analytical approaches, limiting comparability. Such a risk can be mitigated by building servicewide management information systems (MIS) to harmonize and aggregate measurements across government, and by embedding local analytics teams in governmentwide communities of practice. As the case studies in chapter 9 focusing on human resources management information systems (HRMIS) show, integrating different data sources can enhance analytical impacts. Chapter 9 describes the stages in developing harmonized management information systems focused on public administration, and outlines key decision points and trade-offs involved in building public sector data architectures. It warns against constraints in existing legislative institutional environments impeding experimentation with integrating data sources. Such experimentation enables the costs and benefits of integration to be clearly identified, providing important inputs into any scale-up decision. Introducing legislation to allow for small-scale experimentation in measurement and data integration can generate precise inputs and demonstration effects to inform discussions about how to push forward a government's analytical agenda.

Even within existing legislative and institutional constraints, a range of actions can be taken to strengthen analytical integration. A basic activity is the recording of common metadata for all data and existing integration in a centralized repository. This can promote clear communication across government and facilitate public scrutiny. Another action is to monitor what analytics are being used and by whom (see chapter 7). By taking this step, analysts can turn the lens on themselves, and assess how well the architecture of analytics they have developed is holding up, and whether analytics are being used purposefully rather than abused.

## **BUILD AN ANALYTICS AGENDA OF THE FUTURE**

### **Develop Analytical Agendas Everywhere**

Every institution comes with its own staff, tasks, and culture. Thus, the specific analytical requirements of any unit, organization, or public service will vary over a particular task, space, and time. At the same time, for each activity and employee, the questions of what success looks like and how to measure it remain relevant. As such, an agenda for government analytics can be defined at a very granular “micro” level. Every member of the public service can apply an analytical lens of measurement and analysis to their activities, and as such can define an analytical agenda for themselves.

Yet what success looks like and how it can be measured are not central concerns in many government agencies. A first step in resolving this is for actors across government—from senior managers, through unit supervisors, to individual employees—to articulate their commitment to using government analytics where beneficial in their work and to define associated analytics agendas for their areas of responsibility. To institutionalize this approach, performance appraisals could include a compulsory component on the curation of an analytics agenda for an official's areas of responsibility. Organizations could be required to publish an annual update on their analytical agenda. And government could have a central strategy for adopting or expanding government analytics (chapter 26).

None of the discussions in the *Handbook*, or in this chapter, insist that everything in government that can be measured should be measured. Measurement and analysis are costly, and have opportunity costs. Part of an analytics agenda should be to identify an optimal level of analytical investment. Such a level will be iterative, to be updated by the results from previous rounds of analytical investments.

The investment in analytics has potentially high returns. Chapters 12 and 15 show how administrative case data can be used to identify public agencies or individuals who overpay to procure goods for government or take considerably longer than their peers to complete tasks, for instance. Supporting these agencies



or individuals to harmonize their practices with those closer to the average can yield substantial cost savings. Although not all analytics will yield such a return on investment, searching for and prioritizing those that do is a shrewd financial investment. In this vein, evidence from the private sector suggests that data analytics can drive organizational productivity and profitability.

Although the discussion in the *Handbook* has been framed in terms of central government, much of what is discussed applies to any public administrative environment, including subnational public administrations. Subnational entities face some distinct management challenges, such as the management of jurisdictional boundaries. However, subnational analysis can capitalize on the fact that many subnational entities within a country have comparable units with the same functions that can vary considerably in performance and quality. Coordination of analytical agendas across subnational entities has analogous benefits to the centralized analytical units discussed. Institutions that can help coordinate government analytics across subnational entities as part of a community of practice will capitalize on those benefits.

A strength of anchoring analytics agendas in the public administration is that they have a greater chance of continuing beyond any single government administration. By making an analytics agenda a basic part of the administration of government, it can take a longer-term perspective than political agendas can (see chapter 18). This is important because the credibility of measurement and its use for strengthening public service matters for the quality of that measurement. If public servants do not believe that survey results will be used for action and as a management tool, response rates fade, for instance (chapter 26). By clearly signaling that analytics will only be molded but not undermined by political actors, it is likely to be of higher quality. Analytics can build credibility over time, and with stability gain a degree of familiarity.

Similarly, by embedding analytical agendas in public administration, many political leaders are more likely to accept the preexisting analytical architecture as a foundation for their own efforts to strengthen the administration. This may shift political actors toward evidence-based management of the public service.

### **Build Comparable Measurement**

Many elements of government functioning can be usefully measured across sectors and organizational units. For instance, many features of budget, payroll, human resources management, and process quality have commonalities across all of government (chapters 10 to 13). Thus, centralized analytical agendas should push for the harmonization and joint analysis of these features, and agencies should be open to supporting these public goods. Other measures—such as those related to budget utilization and task completion—are more challenging to compare across tasks, but are central to organizational decision-making in areas such as budget allocations (chapter 17). Thus, making explicit the assumptions of such comparisons, and then refining them, is better than skewing measurement toward the most measurable and comparable areas.

Similarly, individual governments should invest in measures that allow international comparisons, such as internationally standardized modules in public servant surveys. Within suitable comparison groups, such harmonization does not substitute for, but powerfully complements, internal measurement. Concerns regarding comparisons of officialdom across sectors or tasks within a single public service can be balanced against concerns regarding comparisons across countries. Having access to measurement from multiple ministries of health around the world will support a health minister's understanding of their own organization's particular strengths and weaknesses in a way that is complementary to their comparison to ministries of agriculture and education in their own countries (chapter 24). To this end, the Global Survey of Public Servants (Fukuyama et al. 2022) aims to increase the volume, quality, and coherence of survey data on public administration over time (see chapter 18). It presents both harmonized data from existing surveys and a suggested common set of questions to be included in surveys of public servants to aid comparability of data from any specific setting. It also recognizes the boundaries of such comparisons and provides access to data (chapter 24).

Comparisons across and within governments can also be made based on administrative data. When such comparisons are made, a frequent challenge is that different organizations in government complete different tasks. One approach the analyst can take is to focus on homogeneous units that do very similar work, such as procurement units across agencies. This is particularly useful for analysts focusing on a specific sector, such

as those described in chapters 14 and 15. As a more general principle, however, such an approach is liable to skew the analytics of government toward areas that are easier to measure (chapter 4). An analyst will gain a more comprehensive picture by defining a holistic agenda for understanding public administration and defining areas where comparability is useful. This relates back to the capacity for public servants to discriminate between when analytics rests on assumptions that fit their setting and when they do not.

With this in mind, when comparing organizations based on administrative data, the analyst should address three questions: (1) Is such a comparison being made implicitly somewhere in government, such as the ministry of finance (for example, when it compares task completion or budget execution rates across organizations)? (2) Can adjustments be made that will make the comparisons more valid (such as by measuring the complexity of the underlying task)? (3) Are there subgroups of comparators for which comparison is more reasonable? As these questions suggest, taking an analytical lens to the issue of comparability itself sometimes allows unspoken assumptions to be surfaced and discussed. Much comparison occurs in public administration without proper care being taken that the underlying issues surrounding comparability are understood and factored into decision-making based on the comparison.

## Use Experimentation Broadly

Faced with uncertainty or ambiguity, how should decision-makers proceed? As tech firms have realized, the optimal choice is to experiment with multiple sensible choices and measure which one works in what environments. For example, according to the company's "rigorous testing" blog, Google "ran over 700,000 experiments that resulted in more than 4,000 improvements to Search" in 2021.<sup>1</sup>

Experimentation in the field of government analytics allows the analyst to trial distinct approaches to measurement of public administration, or the design of government itself, and assess, through proper measurement, the advantages of each. The use of experimentation in the design of public administration is growing in policy circles, and a complementary academic literature is burgeoning in public administration, economics, political science, and beyond. Within this *Handbook*, chapters use experimentation to shed light on methodological questions, such as how the mode of a public servants survey affects responses (chapter 19) and how responses change when questions focus on organizational-level or individual-level referents (chapter 23).

The overlap in work programs across the public sector, and across public sector organizations around the world, presents an opportunity for the repeated testing of successful approaches from other settings, both in measurement and policy. This *Handbook* has illustrated key ideas in government analytics from specific governments. The lessons provide a starting point for undertaking methodological experiments (akin to the A-B testing of large technology firms) in the use of government analytics in a particular setting. In their specific analytics agenda, one question an analyst should aim to answer is: "Does what worked elsewhere work here?" In addition, there is a global benefit to repeated replication of any approach to measurement. Repeated testing of measurement approaches in different settings allows extensions of the findings in this *Handbook* and advances global knowledge toward "stylized facts" about what works where. This will also enhance the quality of conversation on how government functions, grounded in the empirical realities of the service, rather than only perceptions and tacit knowledge.

## PREPARE FOR NOVEL DATA SOURCES

### Search and Testing

The speed at which analytics is evolving requires a constant perspective on novel data sources. A new way of measuring and assessing government functioning could appear at any time. Thus, governments should set

themselves up to capitalize on novel data sources. This requires an approach to analytics that experiments with new approaches to measurement and analysis without the need for wholesale change. Analytics agendas should include an approach to searching for and testing innovations. Individual analysts can assist the search process by publicizing their experiments, by collaborating with others on testing, and by being open to the insights presented by others in the public and private sectors. Centralized analytics units are perhaps the most common way for an organization to engage with new approaches to government analytics.

Setting up an analytics agenda that has a component of search and testing requires a legislative environment that allows public officials a space for experimentation (chapters 9 and 26). Thus, how a government is built will affect its ability to experiment (chapter 16). Institutions that can take the responsibility for the failures that naturally come from testing innovations increase the incentives to experiment. Complementary cultural shifts that reward smart experimentation irrespective of the outcome often require support from senior leadership and political actors. Political actors who can articulate the case for experimentation to their peers and the public buy senior administrative officials space to improve the quality of government administration.

### The Limits of the *Handbook*

There are areas that this *Handbook* has not covered where some governments are starting to make inroads in their analytical efforts. In the area of recruitment, sentiment analysis toward public sector jobs and a wide range of recruitment analytics—for instance, on the diversity of the application pool or the extent of competition for different public sector jobs—can be drawn on by government to improve its quality of personnel. Analysis of communications between government employees, enabled by off-the-shelf solutions from large technology firms, is being experimented with; properly managed and with due care for employee privacy, it promises an understanding of how organizational structure affects team dynamics. Connecting tax records with procurement and customs data can enable an understanding of how government procurement policy affects private businesses and international trade. Machine-learning approaches to images can allow governments to automatically cross-check progress records in infrastructure projects with photos of those infrastructure projects to detect anomalies. And so on.

The *Handbook* has limited the data sources it presents to those of widest use to the largest number of public service organizations. All such entities must deal with payroll and budget, processes, and measures of task completion. Yet this focus on the most standard data sources in government has meant that the *Handbook* has not included some innovative approaches to assessing government functioning.

For example, substantial efforts have been made in geospatial analysis of the impacts of public policy, but there is little evidence that this has been applied to the public administration of the state beyond simple geographic comparisons. Matching personnel with geolocated project data will allow analytics to shed light on whether managers are better or worse at managing projects closer to their homes, or whether there are strong links between characteristics of local labor markets and the quality of recruitment into the public administration in that area. As the world shifts further toward remote work, the utility of tracking exactly where a public official is working and how this affects their productivity may allow for more sophisticated telework policies.

The potential for applying machine learning to text analysis of the vast quantities of documents produced by the public sector is in its infancy (chapter 16). Given that much public service communication is now online, such text analysis and machine learning might be applied to the communications of public officials in real time, and provide automated interventions when there is evidence of a personnel, management, or public policy issue arising.

As governments become more capable of integrating their electronic data systems, the capacity to build maps of the networks of government officials and firms will increase, and it will be easier to assess how personnel who move across different tasks (such as from managing procurement to budget) prosper in different environments and with different colleagues. Overall, gaining a greater sense of what the informal coalitions

in public administration are that facilitate strengthening of government may require triangulation between different data sources.

All these examples underscore the point that a comprehensive analytical agenda is forward-looking, capitalizing on what is available today and readying itself for what might be useful tomorrow.

### The Continuing Validity of the *Handbook*

A number of the foundational themes highlighted in the *Handbook* will continue to be of relevance to any innovations in the field. These include a robust discussion of the ethical implications of government analytics, the boundaries of measurement, and the rigor of analysis.

In terms of ethical issues, the use of data by governments on their own employees has received very little attention, as chapter 6 notes. Although checks and balances exist in public service, these will not always be well-equipped to deal with the pivot to government analytics. Where governments have begun to undertake government analytics, efforts have often not been complemented by a corresponding discussion of the ethical issues involved. For instance, it is important to have robust, servicewide debates about questions such as the extent to which analytics on public officials' remote work communications be undertaken at the level of anonymized individual email or message exchanges, and the ways in which this influences officials' behavior and the capacity to have wide-ranging and honest discussions about public policy.

It is key that such debates are undertaken both sectorwide and within specific organizations because what is considered as ethical and morally right can be very dependent on context (chapter 6). For example, what obligations of transparency around individual activities come with seniority in public service, and how much should officials be actively involved in this debate as they rise up the ranks? Chapter 6 presents a framework for evaluating the ethics of measuring and tracking public sector workers that will continue to be useful to evaluate the impact of innovations in measurement and analysis.

Similarly, the framework presented in chapter 4 will facilitate discussions around the relationship new measurements have to a holistic investigation of the environment being examined. Every new measurement or piece of analysis should come with a "health warning" regarding the boundaries of what it measures, and what it is likely missing. The principles outlined in chapter 5 serve as benchmarks by which new methods can be assessed for their credibility and transparency. Chapter 7 reminds us to turn the analytical lens on analytics themselves and continuously monitor what and how analytics are being (mis)used. And the principles of holistic measurement illustrated in chapter 14 push us to question the extent to which we have "triangulated" any specific measurement with others as a means of capturing distinct dimensions of a variable.

The insights offered in the *Handbook* can strengthen some innovations in measurement and analytics. Better measures of budget or task completion will still rely on the principles outlined in chapters 11 and 17. Innovations focused on improving data quality, availability, regularity, and connectedness will all need to implement the basics outlined in this *Handbook*. Chapters 10, 11, and 12 explicitly discuss layers of data quality that innovations in local settings will help achieve. Similarly, some innovations will build infrastructures that enable more regular, secure, and timely data collection (chapter 9).

## GOVERNMENT ANALYTICS IN AN INCREASINGLY COMPLEX WORLD

As measurement, data, and analysis become the central mediators of decision-making, government must build its capacity to engage with greater complexity in the world and in its own architecture. The question is whether public organizations will reform themselves sufficiently fast so that they can keep up. A solid machinery for government analytics can help empower government organizations to do so.

This chapter lays out the key components of a strategic review process for government actors to think through how they are building a government analytics system that responds not only to today's demands,

but also those of the future. Such thinking is useful at every level of government, from a project manager assessing how they are using administrative diagnostics in their project to the most senior management of the public service thinking through how they might optimally manage service staff.

The lessons presented in this chapter are drawn from across the chapters of the *Handbook*.

The *Handbook's* inability to cover all potential sources of government analytics mirrors the fact that governments will have to prioritize their investments in measurement and analysis. To make those choices strategically, a governmentwide vision of the future, linked to diverse analytical agendas of officials across government, will define the objectives of analytics. Managers who are aware of the trade-offs involved, and supported by specialized offices, will balance investments in basic measurement and the testing of innovations.

As the world gets more complex, the demands on public managers and decision-makers will increase as they manage a more complex government in response. Making the public administration fit-for-purpose will require an informed conversation throughout public service that drives the requisite cultural change. This *Handbook* hopes to inform that conversation. Important public sector conversations regarding reform may occur in a department of local government, a ministry of civil service, or even span countries and the international community. It is therefore important for all government actors to make an informed choice today about how they are setting up a system of analytics that will define what they will know tomorrow.

## HOW TO USE THE HANDBOOK

The chapters in the *Handbook* aim to be freestanding overviews of a particular topic in government analytics and can be read independently. The book is accompanied by a website with annexes and tools for analytics that enable readers to immediately apply insights from the *Handbook* in their own work ([www.worldbank.org/governmentanalytics](http://www.worldbank.org/governmentanalytics)).

To make the best use of the *Handbook*, readers are encouraged to choose the chapters that provide guidance on the data sources most relevant to the management challenges they are facing. For instance, if fiscal sustainability is the core challenge, consider focusing on chapters related to data sources that can yield solutions, such as chapter 10 on the payroll and chapter 11 on budget data. Table 2A.1 at the end of chapter 2 provides a tool to map areas of interest and data sources to the content of the chapters.

The *Handbook* aims at three main external audiences: government analytics practitioners (in governments, international development organizations, and elsewhere); educators; and researchers.

## GOVERNMENT ANALYTICS PRACTITIONERS

The *Handbook* has been designed to make use of the most widespread sources of data on public administration and to address some of the most pressing problems in managing government. As such, our hope is that government analytics practitioners will be able to find inspiration and useful advice in each of the chapters. We also hope that they will see the connections between their immediate interest and other data sources that might enrich the analysis they originally envisaged.

For readers interested in building the analytical capabilities of their organization, this chapter provides a vision of how government might move itself toward being more able to undertake analytics. Chapter 9 describes how to generate an integrated management information system for government. Chapter 26 provides a case study of the US government that presents the complementary management infrastructure that catalyzes any physical data system to become a platform for action.

For readers interested in making the most of their analytics, consider chapter 7 on how to measure whether government analytics are being used and chapter 25 on how to use results from surveys of public servants to strengthen public administration.

For those interested in how different data sources fit together, consider chapter 4 on holistic measurement, and chapter 8, showcasing how analytics can be combined to understand corruption holistically.

Readers looking for practical general statistics tools should go to chapter 5.

For those seeking guidance to think through the underlying ethical considerations of any government analytics effort, turn to chapter 6.

## EDUCATORS

Instructors in a school of public administration or public service training center, or in an academic institution, for instance, should pick and choose areas of particular interest and adapt lessons to the time available.

A single session could provide a useful overview of government analytics. Beginning with the motivation for government analytics (chapter 1), the class could then review a summary of approaches available outlined in chapter 2, and then focus on one particular data source of interest to the use (such as how to use procurement analytics).

A potential module on government analytics could proceed as follows. After an introductory session discussing chapters 1 and 2, consider a class summarizing chapters 4 to 6, to give students a sense of foundational considerations in government analytics. Students could be asked to consider the right ways to apply and manage statistical tools, the ethical considerations particular to studying public administration, and ways to measure holistically in public administration. Perhaps, students could design their own analytics study of public administration that has a pre-analysis and ethics plan that accords to the messages in these chapters.

The third session could focus on chapters 18 and 27, to give students a sense of comparative public administration around the world, and how to diagnose them. The discussion of these chapters could act as an introduction to what data sources are available for government analytics.

Chapter 27 introduces methods for using household surveys to understand public administration, which are the foundations of the World Bank's Worldwide Bureaucracy Indicators. Using the indicators in conjunction with the reading in chapter 27 allows students to understand the global footprint of the public administration, and its relationship to the private sector.<sup>2</sup>

Similarly, chapter 18 outlines the surveys of public administrators undertaken on a regular basis around the world. This chapter complements the data provided by the Global Survey of Public Servants initiative so as to provide the most comprehensive window into the public administration available to date based on surveys of public servants (Fukuyama et al. 2022).

For those students interested in undertaking their own surveys of public officials, the methodological lessons in chapters 19 to 25 provide useful inputs to their design process. These methodological considerations could be covered in a further teaching session on how to do surveys of public servants.

In subsequent sessions, instructors could cover different data sources introduced in parts 3 and 5, focused on the data sources of greatest interest to students. For instance, sessions could cover how to use payroll data, procurement data, and citizen survey data. These sessions should make use of publicly available data sources for students to practice analyzing these data sources.<sup>3</sup>

A teaching module could conclude with a discussion of how to build the analytical capability for government analytics (chapter 3), and how to integrate different analytics sources to assess management challenges holistically (chapter 8).



## RESEARCHERS

Overall, the *Handbook* discusses how to repurpose or construct a range of data sources that are rarely used by scholars, yet provide a fascinating window into public administration and government productivity. For many of the data sources discussed, the *Handbook* is the first consolidated attempt at discussing appropriate measurement. It is one of the goals of the *Handbook* to encourage researchers to expand and improve on the measurement of data sources for government analytics through their work. These researchers—in the fields of public administration, economics, management, political science, or elsewhere—may be in traditional research centers, or from inside government itself, perhaps in an analytics unit focused on improving their part of the public service.

A key early consideration of any research project is what the ethical framework is in which research questions and designs are produced. Chapter 6 provides a useful lens for a researcher to evaluate the ethical implications of their research approach.

Given the weight placed on the rigor and reproducibility of any data analysis, chapter 5 provides a reminder of the principles of good data analysis, and links to a set of resources to make those good practices straightforward to apply. Similarly, given the importance of understanding the limits of interpretation of any single data source or study, chapter 4 provides important reminders as to the validity of any single empirical study or approach.

Part 3 on administrative data can help researchers gain insights into how to construct a broader range of data to better understand the state. Some data sources have featured centrally in recent scholarly work, such as procurement data (chapter 12).<sup>4</sup> Other data sources explored in the *Handbook*—such as payroll data (chapter 10), task completion data (chapter 17), or process data (chapter 13)—have been seldom studied.<sup>5</sup>

Part 4 on survey data presents a range of methodological work related to investigations by the editors and others into how to undertake public servant surveys. Although, as outlined in chapter 18, surveys play an increasingly important part in managing the public sector in a number of countries, rigorous research on how to navigate the decision points that arise in designing, implementing, and interpreting surveys of public servants is limited. Chapter 2 presents a decision tree (figure 2.4) that might be useful to arrange thoughts on factors to be addressed in the survey approach chosen.

Research on the public service is not contingent on having access to proprietary government data. Though some public institutions are making their administrative data publicly available in one form or another, this is the exception rather than the rule. Part 5 presents four approaches that researchers have undertaken to understand features of the public administration using assessments that can be undertaken “external” to the public administration. Each of these data sources can be analyzed by researchers independent of government partnership.

We hope future research on public administration, whether in the fields of public administration, economics, management, political science, or elsewhere, will further capitalize on the data sources outlined in the *Handbook*. With the intention of the *Handbook* evolving in response to new methodological insights in government analytics, we look forward to reading your work or hearing from you.

## NOTES

In formulating this chapter, the authors benefited from discussions with Donna Andrews, Pierre Bachas, Jurgen Blum, Gero Carletto, Verena Fritz, Galileu Kim, Florence Kondylis, Tracey Lane, Arianna Legovini, and Daniel Ortega.

1. See <https://www.google.com/search/howsearchworks/how-search-works/rigorous-testing/>.
2. All of the code associated with chapter 27 is available online. Thus, students can extend the methods presented in the chapter to a country and corresponding household survey of their choice. Such an extension provides an opportunity to work directly with household survey data and learn about what underlies the comparisons made in the indicators, as well as get to study a particular public administration in detail.

3. Payroll data, for instance, are made public by governments such as Brazil (<https://portaldatransparencia.gov.br/servidores/orgao?>). Similarly, citizen survey data are available on topics such as satisfaction with public services (see, for example, <https://www.gu.se/en/quality-government/qog-data/data-downloads/european-quality-of-government-index>).
4. See, for example, Bandiera et al. (2020); Dahlström, Fazekas, and Lewis (2021).
5. There are exceptions. See, for example, Rasul, Rogger, and Williams (2021).

## REFERENCES

- Bandiera, Oriana, Michael Carols Best, Adnan Qadir Khan, and Andrea Prat. 2020. "The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats." NBER Working Paper 26733, National Bureau of Economic Research, Cambridge, MA.
- Dahlström, Carl, Mihály Fazekas, and David E. Lewis. 2021. "Partisan Procurement: Contracting with the United States Federal Government, 2003–2015." *American Journal of Political Science* 65 (3): 652–69.
- Fukuyama, Francis, Daniel Rogger, Zahid Husnain, Katherine Bersch, Dinsha Mistree, Christian Schuster, Kim Sass Mikkelsen, Kerenssa Kay, and Jan-Hinrik Meyer-Sahling. 2022. Global Survey of Public Servants. <https://www.globalsurveyofpublicservants.org/>.
- Google. No date. "Improving Search with Rigorous Testing." <https://www.google.com/search/howsearchworks/how-search-works/rigorous-testing/>.
- Rasul, Imran, Daniel Rogger, and Martin J. Williams. 2021. "Management, Organizational Performance, and Task Clarity: Evidence from Ghana's Civil Service." *Journal of Public Administration Research and Theory* 31 (2, April): 259–77.
- University of Gothenburg. No date. European Quality of Government Index. <https://www.gu.se/en/quality-government/qog-data/data-downloads/european-quality-of-government-index>.