CHAPTER

4

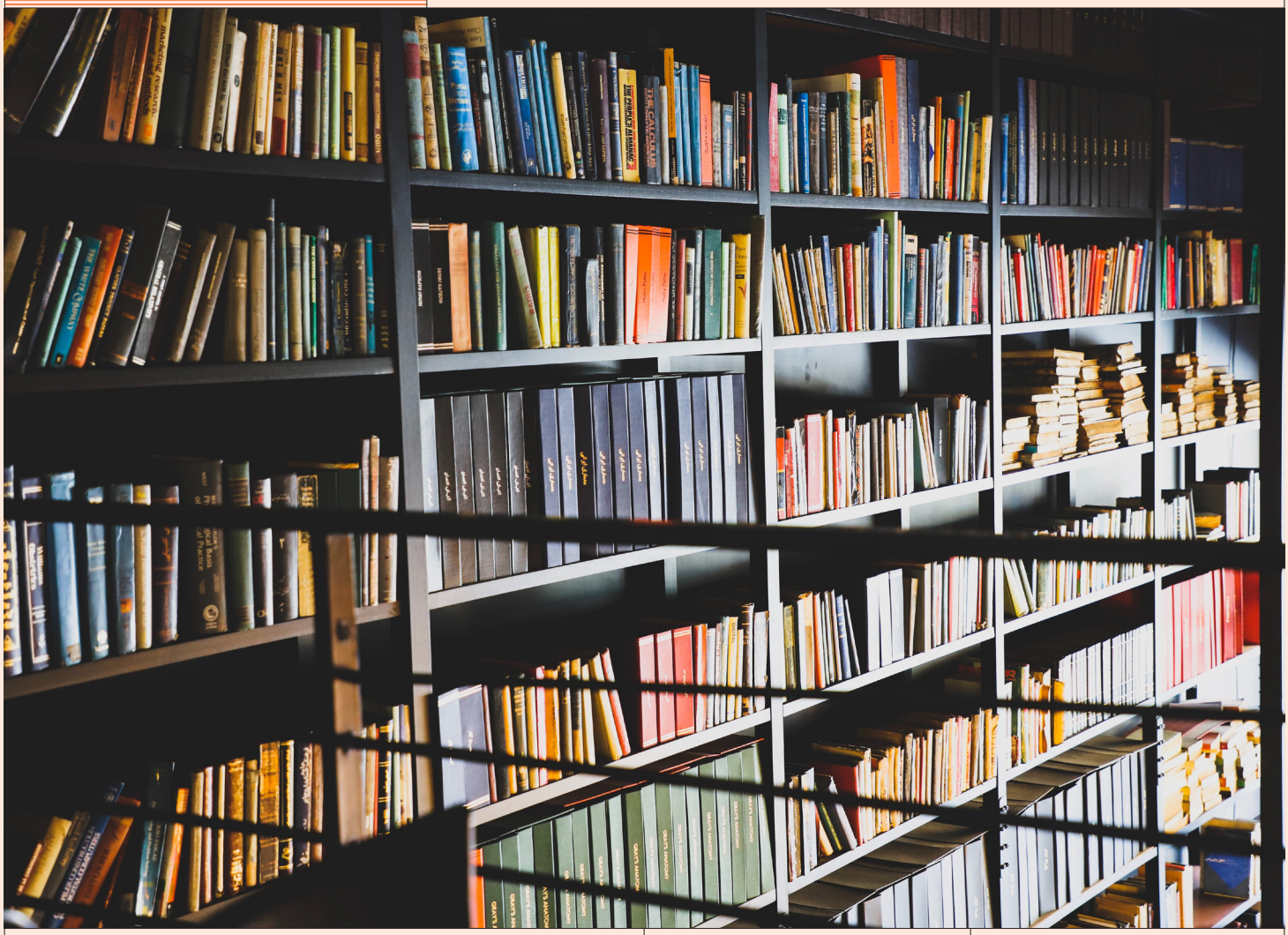
Data for Better Policy Making

4.1

Inadequate data and information systems constrain efforts to improve the quality of spending

4.2

Improving the collection and management of data to support better spending



This chapter is part of the World Bank's 2020 Public Expenditure Review for Indonesia.

This full report is available for download in English and Indonesian via
→ WORLDBANK.ORG/IDPER

To receive the PER and related publications by email, please email
→ YSOEPARDJO@WORLDBANK.ORG

MINISTRY OF FINANCE REPUBLIC OF INDONESIA
Canada
EUROPEAN UNION
Swiss Confederation
Australian Government
THE WORLD BANK IBRD • IDA

4.1

Inadequate data & information systems constrain efforts to improve *the* quality *of* spending

Data are key to measuring and driving effective government performance. Broadly speaking, two types of data are needed to evaluate the quality of spending:

1 Fiscal data on government spending (inputs) classified according to type (economic classification), function, and policy purpose (program/activity); and

2 Sector-specific data on outputs (e.g., the number of schools built, or immunization coverage rate) and outcomes (e.g., student test scores or stunting rate).

These two types of data are necessary to measure the relationship between inputs and outputs (allocative and technical efficiency) and between outputs and outcomes (effectiveness). These data should be available at both the central and subnational levels, and sufficiently disaggregated to undertake meaningful analysis.

A *Data on inputs*

Indonesia has made notable progress in monitoring and reporting spending data at the central government level. Since 2015, the GoI has also fully implemented the electronic State Treasury and Budget System or SPAN (*Sistem Perbendaharaan dan Anggaran Negara*), an automated payment and budget execution information system that provides timely information on the financial position. SPAN is now being used in 222 locations across Indonesia and manages all financial transactions performed by over 24,000 government spending units.[122] The information contained in SPAN enables the MoF and other core financial agencies to produce comprehensive reports on the use of the central government's resources in a timely and accurate manner.

However, the classification of spending makes it difficult to analyze some types of spending in detail. Spending by the central government is regularly reported by economic classification and by standard functions/sub-functions.[123] Indonesia follows international standards (the Classification of the Functions of Government, or COFOG) in the classification of functions at the level of divisions (*fungsi* or functions) and groups (*subfungsi* or sub-functions), but does not use the third level of the functional classification (classes). This makes it more difficult to some types of spending which are of importance to government. For example, some types of infrastructure spending are captured at level 2 of COFOG (water supply, housing, street lighting, waste

TABLE 4.1 Data on subnational spending reported by economic classification and by function, 2014-18

	Spending by economic type			Spending by function			
Year	No of SNGs	Date of data set	Amount (IDR trillion) rounded	No of SNGs	Date of data set	Amount (IDR trillion) rounded	Completeness of function data
2014	542	21-Oct-16	799	324	10-Apr-17	541	68%
2015	542	4-Jul-17	916	529	2-May-17	238	26%
2016	542	18-Oct-18	1003	503	18-Sep-17	667	67%
2017	542	18-Apr-19	1058	542	5-Sep-19	1043	99%
2018	542	5-Sep-19	1092	542	5-Sep-19	1088	100%

Source: Ministry of Finance: http://www.djpk.kemenkeu.go.id/?p=5412

management and waste water management), but others are only captured at level 3, (roads are captured at level 3, under Transport; and irrigation is not separately captured at all, but is a component of spending on Agriculture). Since Indonesia does not use level 3 of COFOG accurately, capturing monitoring infrastructure spending is not straightforward. Furthermore, Indonesia does not classify intergovernmental transfers as spending, as is the international practice under the Government Financial Statistics (GFS) standard issued by the IMF. This is likely to mean that transfers have to be classified by function as part of a manual collation process.[124] Analyzing central government spending on the education function is hampered by the way the largest single expense, salaries, are recorded. In the budget, salaries are shown as a single amount against each Directorate-General, which means the planned costs of delivering individual activities does not include the largest cost item. At the point of execution, salary spending is not captured by sub-function, but instead is classified as 'general government'. Finally, both in budget and spending reports, it is not possible to distinguish spending on religious teachers from spending on religious education administrators, spending on religious teachers in non-religious schools from spending on teachers in religious schools, nor is it possible to distinguish spending by level of the education system.[125]

Data on SNG spending mapped to key functions are available from 2014 to 2018 but are less credible for some functions and for earlier years. Low credibility of data for some functions results from inaccuracies in the mapping of the subnational functional classifications to national ones. Following regulations prescribed by the MoHA, SNGs report their spending according to a more granular set of 34 functions (*urusan*) prescribed in Law No. 23/2014 on Regional Autonomy. A MoHA regulation maps *urusan* to the 11 functions used by central government, but this mapping is not accurate.[126] The MoF has made a significant effort to improve the completeness of spending reported by function, but data are only reliable for 2017 and 2018. Table 4.1 compares data on subnational spending reported by economic classification (left columns) with that reported by function (right columns) on the website of the MoF's Directorate General of Fiscal Balance (*Direktorat Jenderal Perimbangan Keuangan*, DJPK) website for 2014-18 as at December 2019.[127] For earlier years the dataset on spending by function is incomplete as to the number of districts covered, but for 2015 the total reported by function is only around one-quarter of that reported by economic type.

The decision to switch to reporting subnational data according to the 11 national functions has limited the scope to track spending on infrastructure, an important area of spending for the GoI. Prior to 2014, the reporting of subnational spending by *urusan* meant it was possible to estimate subnational spending on infrastructure by combining two *urusan* (public works, and housing and sanitation). Now that subnational spending is reported by nine functions, it is more difficult to identify infrastructure spending. Whereas at central government level most infrastructure spending can be identified by level 2 of the functional classification (sub-function), only the first level functional classification is reported for subnational spending. For 2014, SNGs were responsible for more than 60 percent of total public spending on infrastructure, and it seems likely this has increased during the term of the current administration, which has increased allocation to capital transfers (DAK *Fisik*) and required a minimum of 25 percent of DAU to be allocated to infrastructure. The lack of a way to accurately monitor infrastructure spending is a significant hindrance to the GoI in accurately analyzing the quality of subnational spending.

Evaluating subnational spending efficiency within sectors is even more challenging. The regulations on budget and reporting formats for SNGs do not require them to use the standard classifications for programs and activities, which are important for analyzing the efficiency and effectiveness of spending. A recent World Bank analysis of subnational spending information identified around 15,000 unique program definitions (compared the standard, which provides about 210)[128] and more than 170,000 unique activity definitions (compared with the 1,200 provided in the standard) used by districts in reporting their spending. While it is possible to map around 70 percent of programs to the standard classifications, less than one-quarter of activity definitions can be mapped to the standard. The presence of overlapping definitions means that similar spending could be classified in multiple ways, vastly complicating comparison of spending across districts in order to evaluate its quality.

The MoF attempted to improve the quality of subnational fiscal data through a central automated reporting system, *Sistem Informasi Keuangan Daerah* or SIKD, in 2012. Over the past four years the compliance of districts and the quality

122 These include around 12,000 religious schools.

123 Excluding interest payments and subsidies, data on public expenditures in Indonesia are broken down according to economic classification (personnel, material, capital and social) as well as into 11 functions (General Public Services, Defense, Public Order and Safety, Economic Affairs, Environment, Housing and Communities, Health, Tourism and Culture, Religious Affairs, Education and Social Protection). The World Bank Consolidated Fiscal Database reclassifies these 11 functions into 13 sectors (adding Infrastructure and Agriculture).

124 The central government budget (APBN) separates religion from the 'recreation, culture and religion' function in COFOG. Therefore, at the central government level, Indonesia uses 11 standard functions and 82 sub-functions to classify spending.

125 Although MoF regulation (PMK) 102 of 2018 on Classification of the Budget provides separate sub-classifications for primary and secondary education, the budget and spending reports for the Ministry of Religious Affairs use a composite sub-function classification of 'primary and secondary education'.

126 For example, although MoHA Regulation No. 13/2006 specifies that spending against the urusan of public works corresponds to spending against the central government function of economic affairs, an examination of the sub-urusan level shows it actually maps to three different central government functions: economic affairs, housing and public facilities, and environment and spatial planning.

127 See "*belanja per fungsi*" or spending by function, http://www.djpk.kemenkeu.go.id/?p=5412

128 The standard program and activity descriptions are provided in MoHA Regulation No. 13/2006. However, an amendment in 2007 authorized local governments to customize the classification structure.

TABLE 4.2 Comparison of national and subnational functional classification in education under proposed subnational classification system in MoHA Regulation No. 90/2019

Education is a function defined in PMK 102/2018 for central government	Education is a function defined in MOHA 90/2019 for subnational government	
Level 2 function definitions for central government (PMK 102/2018)	**Level 2 function defintions for subnational government (MOHA 90/2019)**	**Level 3 of Education sub-function defined in MOHA 90/2019**
Early childhood education programs	Education	Education management
Basic Education	Youth and Sports	Curriculum development
Intermediate education	Library	Teachers and teaching personnel
Non-formal and Informal Education		Education licensing
Official Education		Language and literature
Higher education		
Educational Assistance Services		
Religious Education		
Education and Culture Research and Development		
Youth and Sports Coaching		
Cultural Development		
Other Education		

Note: Under the MoHA regulation the functional and program classifications are linked. Level 3 of the functional classification shown in the right column is part of the program segment.
Source: MoF Regulation No. 102/2018 for national function classification, MoHA Regulation No. 90/2019 for subnational function classification.

of data has improved substantially, but production of meaningful data from the SIKD system depends on use of a more standard classification by SNGs, which will entail major change management of local accounting and reporting practices. Traditionally, the MoHA has regulated the classification system used by SNGs. Implementation of a more standard approach will require support of other ministries including the MoHA. In addition, the MoF continues to extract data manually from paper reports for the purpose of public reporting of subnational spending. Data are available by economic classification and by function (as shown in Table 4.1 above), but not the intersection of both. Hence, it is not possible to evaluate effectiveness of subnational sectoral spending by looking at the relevant spending mix (e.g., how much do SNGs spend on salaries, capital, and goods and services in the health sector).[129]

A new MoHA regulation on classification of subnational budgets and spending contains improvements but will make it more difficult to obtain a comprehensive picture of total government spending. MoHA Regulation No. 90/2019 provides for additional segments in the subnational budget classification and standardizes the way programs and activities are captured. However, it also fully aligns the classification of programs and activities to the *urusan* classification structure at three levels, which will make it more difficult to consolidate central and subnational spending. A new segment on function is introduced, which uses the national functional classification at level 1, but creates an entirely new classification structure at level 2, as shown in Table 4.2 for the Education function. Whereas a breakdown of spending by level of the education system is possible from the central government classification structure, this will not be possible from the subnational classification structure. In some cases, given the differences between the sub-functional components of each different system of functional classification, the types of spending captured at subnational level will be quite different from that captured at national level.[130] If the GoI wants to analyze total government spending in a rigorous way, it is important that the two classification systems properly align in terms of detail, not just in name.

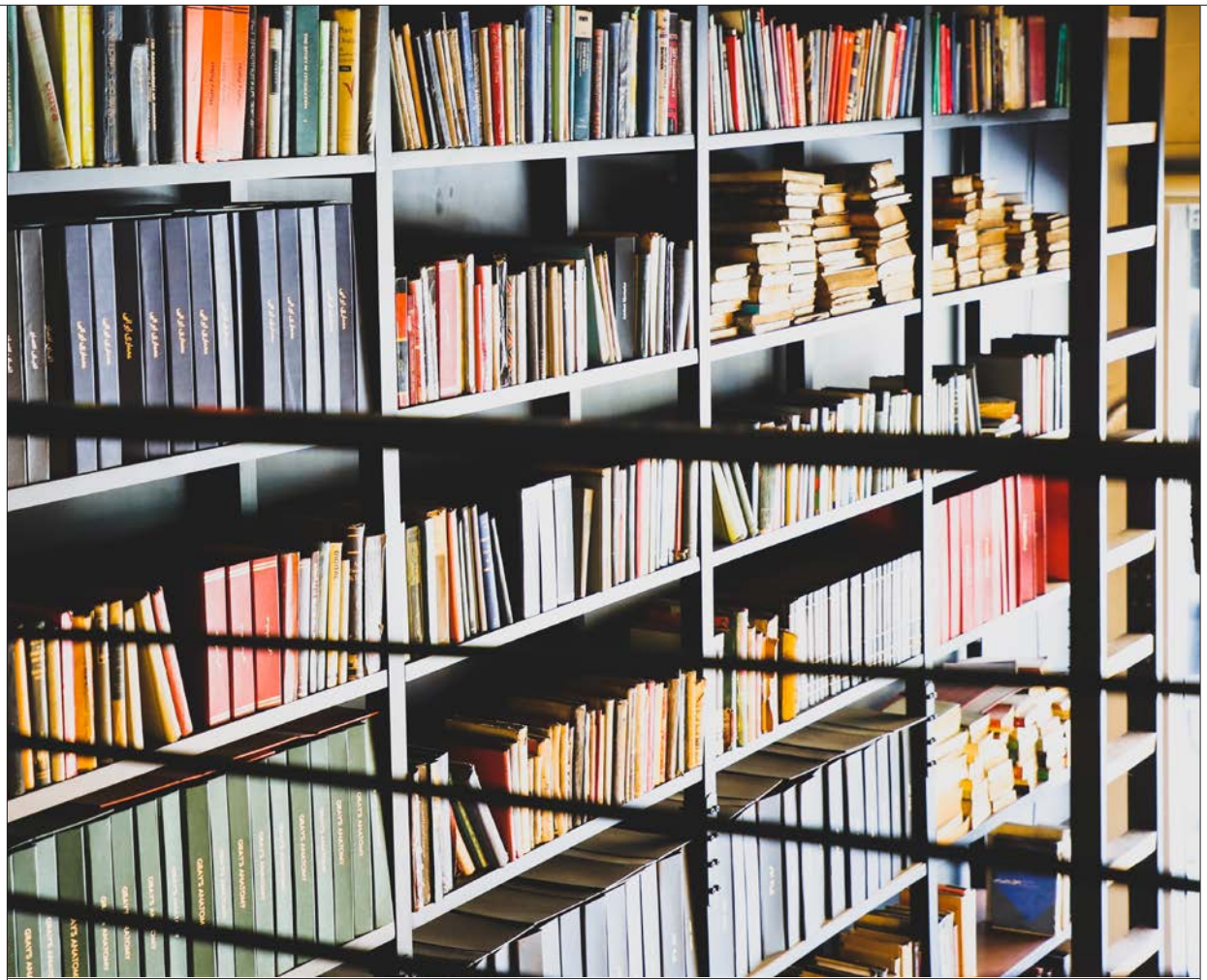
129 For the 2018 budget data, the MoF (DJPK) has published a breakdown of spending on each function into broad economic categories—salaries, goods and services and capital. These data are not yet available for spending, but it is a promising start.

130 For example, although the *urusan* classification of 'public works' (*pekerjaan umum*) at subnational level is mapped to the national function of 'economic affairs', it covers sub-functions which at the national level are mapped to different functions: (i) Solid waste and Waste water, which at national level are classified under the function of 'Environment and Spatial Planning', (ii) Housing and Street lighting, which at national level are classified under the function of 'Housing and public facilities'.

B

Outputs

Data on outputs are available in some sectors but are not consistently used and lacking in quality. Outputs are usually collected through administrative systems maintained by each line ministry. The MoEC has developed a ministry-wide system, Dapodik, an effort that other ministries could emulate. In other sectors, data are highly fragmented across multiple departments of the same ministry and/or prone to different definitions and lack of quality assurance in the collection process (see Box 1 in the chapter on Health and the 2013 report on maternal mortality[131]). Information on the current quality of infrastructure (used to inform a needs-based allocation of capital funding and to measure achieved performance of programs or projects) is captured in similar ways through administrative systems. Such administrative data are prone to manipulation and gaming. If indicators are increasingly used to reward performing SNGs and to name and shame laggards, SNGs will face growing incentives to overreport their achievements or to focus on "hitting the target", while missing the point. A World



Bank-financed project, the Local Governance and Decentralization Project, helped Indonesia to pioneer the use of independent verification to check the validity of self-reported performance assessments for individual subnational infrastructure projects.

A recent IMF/World Bank assessment of public investment management systems has identified a gap in data on public investment projects.[132] In many countries, budget classification systems include a project segment which allows expenditure on capital projects to be monitored more closely during budget execution and tracked across years. The absence of project-level information for tracking capital projects in plans and budgets undermines good management to ensure full budget absorption and efficiency. Given the importance of infrastructure investments for government, this is a major gap.

While data on outputs may generally be reliable for measuring performance at aggregate national level, their use to measure performance of individual districts is more problematic. In the health sector, for example, it is not uncommon for district immunization rates to be well over 100 percent. These errors likely result from inaccurate calculation of the denominator—the number of children who should receive vaccinations (i.e., those born in the past 12 months). Accuracy in measuring outputs at subnational level is not just important for comparing the performance of districts with each other; it is also important to guide district managers where they need to focus attention. This includes information about performance across a single district. For example, current systems for monitoring stunting are designed to produce a robust result at the district level, but they are not reliable for identifying locations where stunting rates are higher within a district.

More generally, there are competing sources of population data of beneficiary target groups, which allows administrative data on outputs to be converted into comparable performance measures. There are two sources of population data in Indonesia: Intercensal and Census surveys (conducted every five and ten years, respectively), and civil registration data collected by the MoHA. Since 2013, a law on civil registration[133] has directed public agencies to use civil registration data (MoHA population data) in calculating entitlements and allocating resources. However, population estimates generated based on the Indonesia Intercensal Population Survey tend to differ starkly from administrative population data as reported to the MoHA. In 2015, the difference in population estimates exceeded 10 percent for over one-third of districts and exceeded 20 percent for about 11 percent of districts.

The GoI is making efforts to improve the quality and coverage of civil registration data. Beyond expanding coverage and underpinning the reliability and sustainability of the national ID system, the quality of demographic and health statistics depends on accurate and timely registration of births and deaths. One reason may be that SNGs only capture those births and deaths that are reported to a Posyandu or a Puskesmas. Birth registration and national IDs also have important implications for removing barriers to the poor accessing health and education services. Increasing access to these data by all ministries and local governments is therefore critical.

131 Joint Committee on Reducing Maternal and Neonatal Mortality, National Academy of Sciences, 2013 Reducing Maternal and Neonatal Mortality in Indonesia, Saving Lives, Saving the Future, Chapter 2 The Data Conundrum. http://staff.ui.ac.id/system/files/users/tjahyono.gondhowiardjo/publication/saving_lives_saving_future.pdf

132 Indonesia Public Investment Management Assessment. IMF, World Bank, 2019. See box in Overview chapter.

133 Law No. 24/2013, Article 58.

C

Outcomes

Data on outcome indicators is usually obtained from the annual household survey, Susenas, or from periodic sector-specific surveys such as Risfaskes, the health facility survey. Survey data provide a more accurate measure of access to services and outcomes but may not be reliable for measuring year-on-year changes at the level of individual districts. Special surveys are often undertaken only every few years, while the routine surveys such as Susenas use a sampling approach, which is not designed to generate a robust result at the district level. For more than 200 districts, the confidence interval for Susenas at the level of individual districts is greater than 5 percent. Since expected year-on-year performance improvements are often much less than 5 percent, Susenas year-on year changes are not a meaningful way to measure districts' incremental performance improvements. Part of the problem is that the sample size for specific subpopulations, such as households with children under five, is insufficient in some, especially small districts. Measurements from Susenas related to infrastructure (such as access to water and sanitation) are prone to additional clustering errors, arising from the way the survey is administered in blocks of 10 households. Use of a rolling average of measurements from annual Susenas surveys can increase the reliability of year-on-year measurement of performance changes.[134]

Where data are available, the lack of better integrated monitoring systems is clearly impeding the GoI's ability to spend better. In the health sector, for example, multiple monitoring systems are managed by different directorates within the MoH for different health interventions, and there are multiple systems to process JKN claims under BPJS Healthcare. With the lack of interoperability between different data systems and poor coordination among key stakeholders, there is limited useful information that can inform strategic prioritization and resource allocation at the district and national levels. Despite improved coordination in the allocation of DAK, decisions on how much to allocate to each district are still based on information from the districts themselves. It is difficult to assess if district proposals are based on a consistent measurement of needs. The introduction of the unified poverty targeting database (*Basis Data Terpadu* or BDT), in 2011, currently known as integrated social welfare database (*Data Terpadu Kesejahteraan Sosial* or DTKS), was followed by a more efficient allocation of social assistance benefits in subsequent years. However, DTKS has not been systematically updated since 2015, and is not fully used by all major social assistance programs. As a result, it has not been able to foster convergence across social assistance programs, i.e., ensure that eligible families receive an integrated network of support.

Without well-functioning information systems, systematic monitoring and evaluation of how public resources are spent will remain challenging. The lack of M&E is evident across all sectors, but particularly in infrastructure, which the GoI has prioritized in recent years. In roads, poor collection of data on asset preservation and development has contributed to fragmented, ineffective prioritization of programs to improve road performance. Although more modern planning tools are starting to be utilized, many *Balai* (regional support teams) still undertake manual screening of pavement conditions using spreadsheets. In the housing sector, the lack of data on the quality of subsidized housing during audits means that there is no mechanism to hold developers accountable. There is also no system in place to systematize and enforce compliance with construction regulations for subsidized housing.

134 B. Lewis, N McCulloch and A. Sacks. 2015. 'Measuring Local Government Service Delivery Performance: Challenges and (Partial) Solutions in Indonesia'. Journal of International Development.

4.2

Improving *the* collection & management *of* data to support better spending

- A *Inputs*
- B *Outputs*
- C *Outcomes*

Improved data are essential to make sure that each rupiah of public money is spent efficiently and effectively in Indonesia. To identify which programs/interventions are working and to undertake evidence-based policymaking more broadly, the GoI needs better data. As previously noted, data on inputs, outcomes and outcomes are often unavailable, not updated regularly or sitting in different systems that are not integrated with each other. The problems are more severe at the subnational level and adversely affect SNGs' ability to deliver better access to services.

A Inputs

Although data on public expenditures by the central government are good by international standards, the GoI needs to ensure that monitoring systems collect the necessary information to drive better performance. As previously noted, the data on spending by the central government are regularly monitored, reported and available to the public. However, Indonesia needs to monitor spending on infrastructure closely, and this analysis is not well supported by the use of functional classifications. In the international standard functional classifications, detail on infrastructure is provided at the third level of the classification, which is not used in Indonesia. In the absence of third level functional classifications, accurate monitoring of infrastructure spending will require a combination of functional and economic classifications. To ensure capital investments are properly managed, a classification for project ID should be introduced. The recent IMF/World Bank public investment management assessment recommended that information on major capital projects should be included in the next RPJMN with information on timeframe and estimated costs (see box in Overview chapter). In order to monitor implementation of planned projects, IT systems such as SPAN should be modified to include a project ID.

Better definition of programs and activities (sub-programs) in the budget classification and Chart of Accounts would support more effective monitoring of interventions. Tracking performance effectively starts with a clear logic as to how the desired outcome will be achieved. In many cases, the delivery of interventions depends on inputs from multiple levels of government. The GoI plans to introduce more consistent classification of programs and activities across levels of government, and to better integrate allocations to national ministries with subnational transfers, which will support better monitoring of the overall envelope for delivery of government programs. Budget classifications could be better aligned with intervention logic and with the priorities expressed in the national plan. As currently structured, program and activity classifications are hardwired to the organization structure, which inhibits meaningful monitoring of performance.[135] The Annual Plan uses a different architecture of classifications from the budget, which makes it difficult to track the links between the two. Further refinement and rigor in the definition of outputs would establish a clearer results chain from inputs to outcomes. Similarly, improving the capture of large infrastructure projects in planning and budget management systems (e.g., in SPAN) would make it easier to track their implementation. One option that could be explored is to require ministries to identify all projects over a certain size as a standalone output in the budget.

Linking SPAN and the procurement system would generate useful data to support expenditure analysis. Currently, the procurement system (SPSE) managed by LKPP focuses on sourcing, whereas SPAN managed by the MoF focuses on recording commitments and payments of the goods and services procured or sourced. Sourcing information from SPSE is not visible in SPAN, while commitment and payment management information from SPAN is not visible in SPSE. Establishing a link between the two systems would enhance transparency, efficiency, predictability and control over budget execution. For example, the GoI could monitor transparency in procurement by looking at the share of contracts that are open to competition. The GoI could also measure the time taken in procurement processes (disaggregated by procurement methods), whether the same vendor gets selected by 'single source' or other non-competitive methods, and whether payments are released at a faster rate in non-competitive contracts. The first indicator would enhance the efficiency of spending, whereas the latter two indicators could be used as a red flag in monitoring corruption.

At the subnational level, recent reforms to improve the quality of spending data are in the right direction but implementing them is a huge task. Initiatives to implement a standard budget classification and Chart of Accounts (*Bagan Akun Standar*, BAS) are underway. Government Regulation No. 12/2019 (issued in January 2019) requires SNGs to budget and report using a common classification system and specifies that a separate government regulation will determine the classification system. The MoF is leading the development of that regulation to define the architecture and definitions of the classifications that SNGs will be required to use. In the meantime, the issue of a separate ministerial regulation on budget classification and Chart of Accounts[136] by the MoHA and the introduction of a new system for managing subnational finances presents a coordination challenge. It will be critical for the MoHA and the MoF to work together to arrive at a harmonized classification struc-

135 Programs correspond to Directorates-General and Activities to Directorates.

136 MoHA Regulation No. 90/2019 was issued in November 2019 and specifies the new system will apply from January 1, 2020. Its implementation is reinforced by the roll out of a new e-planning and budgeting system, the SIPD.

ture that addresses the information needs of each organization, but which also prioritizes the production of meaningful budget and spending reports that support decision-making by SNGs. The level of granularity in subnational plans, budgets and financial reports undermines good accountability.[137] Classification systems should be structured to support good subnational budgeting and budget execution decision-making focused on three objectives: (i) prioritizing across sectors and services; (ii) transparency of allocations across major expenditure types and between frontline service delivery and back-office administration; and (iii) transparency of capital investments. The introduction of a new classification system offers the potential to vastly improve the tracking of capital investment projects at the subnational level, but this would require introduction of a project ID, which is not currently part of the proposals put forward by either the MoF or the MoHA.

The integrity of these important reforms to standardize the classification of subnational spending will depend in high level inter-agency coordination and willingness to evolve the system over time. The task of rolling out the new classification system in 500+ SNGs will be a huge one. At a minimum, local governments will need to map their current BAS to the new BAS, clean the data for transfer to the new system, and maintain audit files on how they have managed the transition process (to meet the requirements of BPK, the state audit agency). It is inevitable that it will take some time to train local government officials in how to apply the classification consistently, and the classification structure will need to be revised as gaps are identified. Other large decentralized countries (e.g., South Africa, Mexico and Brazil) have taken 8 to 10 years to implement similar reforms. To ensure this reform is managed properly, adequate resources should be allocated for dedicated staff to manage the process, and to finance technical support to the 500+ SNGs to collect and classify spending information accurately.

B

Outputs

Data on access, outputs and beneficiaries should be integrated into common platforms and more attention paid to their maintenance. The experience with the integrated social welfare database, DTKS, shows that a well-functioning data registry that is accessible by all stakeholders can yield crucial gains in efficiency and effectiveness. Continuing to update and ensure full implementation of the DTKS would help to improve the impact of social assistance programs on welfare. The MoEC has established and is continuing to refine its Dapodik database, which provides a platform of information on the status of schools under the MoEC. It could be expanded to include religious schools supervised by the MoRA. Meanwhile, other sectors need to take the first step in establishing a common database. In housing, for example, an integrated Housing and Real Estate Information System (HREIS) containing data on key metrics (e.g., housing backlog, substandard housing, and affordability) by geography and consumer income could help policymakers identify gaps between housing supply and demand. In health, a common dashboard to benchmark performance among districts and facilities, available to all stakeholders across levels of government, could be established. Moreover, JKN claims data can help monitor adherence to guidelines and protocol-based care, thus helping improve the quality of service delivery. Claims data could also be used to run simulation and budget impact analyses to help identify cost-savings from open-ended payments to hospitals.

Assessment of relative infrastructure gaps (for example, across districts) is an important component of the central government's redistributive function. Allocation of DAK could be more efficient if it is targeted to jurisdictions with the greatest need, but that would require a more consistent way of measuring need. Minimum standards were intended to serve that function, but the latest refinement to minimum standards has focused more on measuring the services received by citizens rather than the gaps in inputs such as schools, health centers, water supply systems and roads. Some countries use minimum standards specifically for infrastructure, and these could be adopted for Indonesia.[138] Service accreditation systems like that for health facilities could also be used as basis for fair comparison of the relative needs of different districts. To properly inform allocation of capital funding to bring infrastructure gaps, the standards need to provide not just a benchmark for the quality of individual infrastructure assets, but a benchmark for infrastructure quantity as well.

The GoI has already laid a solid foundation to improve the quality of data through the One Data initiative and the recent Presidential Regulation on e-Government. The recently issued Presidential Regulation on One Data (Presidential Regulation No. 39/2019) sets out a whole-of-government approach to data governance to improve government data quality, management and integration across government. In addition to enabling sharing of data within government, this is also expected to improve the transparency, accountability and accessibility of government data for the public. The regulation establishes governance arrangements and standards for data management, covering both central and subnational lev-

137 It is not uncommon for subnational budgets to be over 500 pages long and for individual department workplans to be several hundred pages long. These are prepared and approved annually and routinely revised halfway through the year, resulting in a large transaction burden on local governments which distracts them from better strategic management of good quality spending.

138 An example is the Regulations for Norms and Standards for Public School Infrastructure, issued under the South African Schools Act 84 of 1996.

els. Implementation of the initiative is led by Bappenas, together with MoABR, Ministry of Communication and Information Technology, MoHA, MoF, BPS and Geospatial Information Agency (*Badan Informasi Geospasial*, BIG) on the Steering Committee. Accordingly, the One Data Secretariat will be housed in Bappenas to harmonize relevant policies on data standardization, management and exchange, and coordinate the One Data Forum, while each ministry is expected to appoint a "data custodian" to implement the policies and standards. Government data covered by the regulation not only include statistical data and geospatial data, whose standards are governed by BPS and BIG respectively, but also various data generated as by-product of government administration, such as fiscal data. One of the functions of the One Data Forum will be to establish Master Data and Reference Codes to be used across government which, along with use of common data and metadata standards, as well as requirement to store data in open and machine-readable formats, will be important for enabling data interoperability. This regulation is complementary to the e-Government regulation (Presidential Regulation No. 95/2018) issued in 2018, which focuses on establishing common standards for technical infrastructure, such as Government Data Centers and shared applications systems.

To support the implementation of data improvement with integrity, more attention is needed on the enabling environment for ministries to discharge their data stewardship functions: (i) the capability and financing of ministry data centers (typically housed in Secretary General's Office); (ii) cyber security and information privacy policies; (iii) incentives for civil servants to specialize in data and technology; and (iv) improving the quality of government IT procurement (for example, modelling the UK Government Digital Service function in the Cabinet office, which provides oversight of the quality of IT development for the Government of the United Kingdom).

BPKP (the internal audit agency) has developed skills in verification, and more use could be made of its considerable capacity. Administrative data should be verified, particularly where they are being relied upon to calculate performance incentives. The Local Governance and Decentralization Project supported BPKP to undertake verification of individual DAK-funded projects in roads, water, sanitation and irrigation against a set of standard criteria. BPKP has been appointed as the independent verification agent for World Bank programs for results, of which Indonesia now has four. BPKP has a wide presence across Indonesia and considerable professional capacity, as most of its staff are accountants. There is considerable potential to make more use of BPKP in monitoring. The state audit agency, BPK, has also expressed interest in undertaking performance audits which, beyond ensuring accountability for public resources, could look at value for money in terms of program design, effectiveness of eligibility and allocation criteria in terms of targeting and overall program management effectiveness.

More may need to be allocated to the function of M&E of government programs. While there is understandable caution about allocating resources to costs that do not translate into services or assets, under-spending on M&E is a false economy. Closer examination of M&E systems could yield evidence to make the business case to support increased allocation. Increased funding will be needed to support BPKP's ongoing involvement in monitoring public programs, as well as ensuring data systems are adequately resourced. International practice suggests a rule of thumb of around 10 percent of program cost, higher if the program is executed at community level or involves very significant resources.[139]

Efforts to standardize and verify population data, some of which are underway, should be encouraged and prioritized. The level of under-registration varies markedly from one district to another, even where they face similar logistic challenges. A more targeted combination of incentives and support is needed to stimulate districts which are lagging, reward those which are performing well, and foster innovation and dissemination of ideas on how to improve registration systems. To facilitate improvements in population administration services, the central government has also started providing special grants (DAK Adminduk) to local governments since 2017. The current allocation formula for districts is uniformly based on population, but changes are under consideration to link the allocation and disbursement of these grants to the performance management framework for local civil registration offices (*Dinas Dukcapil*). For lagging regions, the push will be to expand *access* to services and close coverage gaps, e.g., birth certificate coverage, while for the best performing regions, the results focus may shift to *quality* of services, e.g., timely birth registration, compliance with service level standards. More transparency of the discrepancies between different population data sources could help stimulate further improvement. At present, data on civil registration are intended to be published every six months, but up-to-date and complete data by province and district are still difficult to access publicly.

139 F. Twersky and A. Arbreton, 2014, Benchmarks for spending on evaluation. For federally funded community level programs, an allocation of 13% of budget for evaluation is recommended: https://www.nationalservice.gov/sites/default/files/resource/Budgeting%20for%20Evaluation_090914st10.17.pdf. The Treasury of Western Australia recommends that organizations implementing high risk government programs should quarantine between 5-10% of their program budget for evaluation. High risk programs are those involving more than AUD 5 million, which are innovative, or which are a high priority for the government: https://www.treasury.wa.gov.au/uploadedFiles/Treasury/Program_Evaluation/evaluation_guide.pdf. See also 'State of Evaluation 2012' http://www.pointk.org/resources/files/innonet-state-of-evaluation-2012. pdf?fbclid=IwAR2jx0YYwv-VYCSYcoCT4slFOFmI_ZOjBJ2rv1Pql6t2JaP6M_WMbxUfIbI.

C

Outcomes

Indonesia already has one of the most regular and accurate national poverty surveys in the world. There is a risk of over-burdening Susenas and compromising its core function to monitor poverty reduction if more indicators are added to serve supplementary purposes. However, an independent survey capacity is needed to monitor some key outcomes such as reduction in stunting. Producing outcome level information that is accurate at the level of individual districts is challenging and expensive. BPS is a critical agency responsible for producing key information about the social and economic conditions of Indonesia, ranging from economic growth, poverty, employment to prices. With rapid decentralization, urbanization and increasing complexity of the economy, demand for more disaggregated and timely data has increased, even as data collection challenges have also increased. The scope of institutional reforms to continue improving relevance and quality of data to support policy making include: (i) standardization of business processes and statistical infrastructure (to conduct different surveys and/or Censuses); and (ii) application of technology solutions for data collection, management and dissemination.

Recent reforms to computerize test scores are an example of improvements in the reliable measurement of education outcomes. The introduction of computer-based testing for 9th and 12th grade student exams has reduced opportunities for corruption (gaming) in the scores themselves, but the measurement occurs only at the end of the student's completion of the junior and senior school cycles. Taking an outcome measurement earlier, for example when students pass from basic to junior high school, would provide a better opportunity to identify where in the education system challenges are most pronounced, and ensure that students who are not ready for the next stages of school are either given additional support or are not promoted to the next grade.

Stimulating an enabling environment for better data quality

Demand for better data is unlikely to increase unless the data are used. This is particularly true for subnational data. In many sectors, having access to central government spending data does not inform expenditure analysis in any meaningful way, without access to subnational spending data so that there is a complete picture of resources to align with outputs and outcomes. Some actors such as the Ministry of Health have a considerable appetite for expenditure analysis but lack the data to undertake this. Once agencies have access to and are using data, they are more likely to identify its shortcomings and prioritize its improvement.

The budget process is an important entry point to increase the use of data. Ministries should be required to substantiate requests for funding increases, or to introduce new programs, with business cases based on evidence. Periodic spending reviews of major spending programs should also be conducted. Where data are fragmented across sectors (e.g., health, education or infrastructure) due to multiple ministries or stakeholders, annual sector reviews of performance and expenditure should also be required. Rather than attempting systematic evaluation of all government spending, a few programs involving high spending or high priorities for government could be selected to develop and refine more sophisticated qualitative approaches to using spending reviews in the Indonesian context. Spending reviews help to promote ministry accountability for performance, not just accountability for spending. The annual performance and expenditure reviews of the National Stunting Reduction Acceleration Strategy are a good example that could be further expanded to other spending programs.

Integrating systems can promote harmonization of data, but the devil is in the details. Some systems are being established in which data are being collected through PDF uploads, rather than through entry into the system itself. Real integration only comes with the use of a common data structure and closed menus to classify key data attributes to ensure comparability.

Transparency can be a powerful driver for data improvement. To reinforce the accuracy of these systems, key data should be made public. High-level political commitment to the principles of open data could have a catalytic effect on improvement of data quality. *Satu Data Indonesia* or the One Data Initiative,[140] spearheaded by the President's office and Bappenas, is a good start. An expanded One Data Initiative could focus on: (i) improving the integration of data collection, quality assurance and management across ministries; (ii) establishing data quality standards; and (iii) facilitating inter-agency agreement on data exchange. Verification is an important mechanism to ensure that data quality remains consistent. Enabling Parliament, local governments and citizens to access and utilize the data would improve both bottom-up and top-down accountability.

140 https://data.go.id/