
Rethinking the Performance Monitoring and Evaluation System of Provincial Administrations in Armenia

April 7, 2020

By Anush Hambaryan

Harvard Kennedy School of Government

Master in Public Policy Candidate

Expected Graduation: May 28, 2020

Policy Analysis Exercise (PAE)

Client: The Department of Territorial Administration at the Ministry of Territorial Administration and Infrastructure of Armenia

Advisor: Professor Jay Rosengard

Seminar Leader: Professor Janina Matuszeski

This PAE reflects the views of the author(s) and should not be viewed as representing the views of the PAE's external client(s), nor those of Harvard University or any of its faculty.

Table of Contents

<i>Executive summary</i>	4
<i>Background information</i>	8
<i>Territorial administration system of Armenia</i>	10
<i>The existing performance evaluation framework of provincial administrations</i>	13
<i>Policy Issue</i>	14
<i>Research questions</i>	15
<i>Results based monitoring and evaluation systems</i>	18
<i>Analysis of the existing performance monitoring and evaluation system</i>	20
<i>Examples of performance monitoring and evaluation systems</i>	36
North Carolina Benchmarking Project, USA.....	36
County of San Mateo, California, USA.....	38
Australian government, Productivity commission: Report on Government Services	40
<i>Recommendations</i>	43
<i>Annexes</i>	47
<i>Endnotes</i>	49

Policy Question

The Department of Territorial Administration at the Ministry of Territorial Administration and Infrastructure of Armenia wishes to improve the performance evaluation system of provincial governments in order to hold them accountable, and to incentivize better performance. They would also be open to understanding how evaluation could help them learn about programs and inform decision making.

Framework

Evaluation can help with accountability and boost performance, if it tracks the results that provincial administrations are expected to deliver. Results could be in the form of short-term or long-term outcomes. While the latter are usually hard to measure, ideally both should be measured. A general evaluation framework (log frame) would include goals, activities, inputs, outputs, short- and long-term outcomes, as well as indicators for each of these and targets for goals.¹ Governments also set and measure process-related goals, such as compliance with law, efficiency, equity. In addition to these, it is often useful to implement qualitative evaluation in order to better understand the outcomes (for example, why certain programs work, and others don't) and how they can be improved.

Current system in Armenia

While components of the log frame mentioned above (goals, targets, activities, etc.) can be found in programmatic documents (development strategies, annual activity plans), the evaluation system mostly includes only activities, which are graded based on loose instructions with significant leeway for subjective interpretation. The programs in annual activity plans of provincial administrations are not tied to the results expected of them and the evaluation system does not have a comprehensive framework of indicators that would allow for measuring efficiency, effectiveness, equity and other important aspects of programs.

That being said, there are a few areas, where actual indicators are used for evaluation. For example, in finance, compliance with procurement legislation is measured by the share of procurement reports approved by the Ministry of Finance. In Health Care, death and disease trends are part of the evaluation indicators.

Finally, the current system does not cater to learning and no qualitative evaluation is implemented to learn more about programs.

An example of a sector

While there are some differences, evaluation is quite similar across sectors, where provincial administrations have competences. Health care is a typical example, which is described below.

The guideline developed by the Ministry of Health Care specifies 17 sub-indicators for the 4 indicators. Despite being called “indicators”, most sub-indicators are “activities” that provincial administrations have to implement. For example, participation in the streamlining projects of health care system; supporting the implementation of sanitary-hygienic, anti-epidemic and

quarantine activities; maintenance and technical equipment of the buildings of provincial health care institutions, etc.

Some of the activities include no outcomes (for example, organizing the work of provincial health care organizations). Others have outcomes (for example, taking actions aimed at improving the quality of services in primary health care institutions in the province), however the indicators are not meant to measure the outcomes themselves (in this example, the quality of health care services is not evaluated, but the actions taken).

The grading is based on quality of implementation and timeliness. The criteria of quality are the following: 1) Implemented without flaws, 2) Implemented with non-essential flaws, 3) Implemented with essential flaws, 4) Not implemented. The criteria of timeliness are the following: 1) Timely implementation, 2) Non-essential violation of deadlines, 3) Essential violation of deadlines. The guidelines do not specify how the quality of implementation will be assessed, in particular, what constitutes flawless implementation, how flaws will be classified as essential or non-essential. Similarly, it does not specify how violations of deadlines would be classified as essential or non-essential.

Main Recommendations

- Create result chains for the programs outlined in annual activity plans of strategic development documents, where goals, inputs, activities, outputs and outcomes would be logically connected to each other. Make sure that outputs and outcomes are not worded as activities and are in fact output and outcome statements.

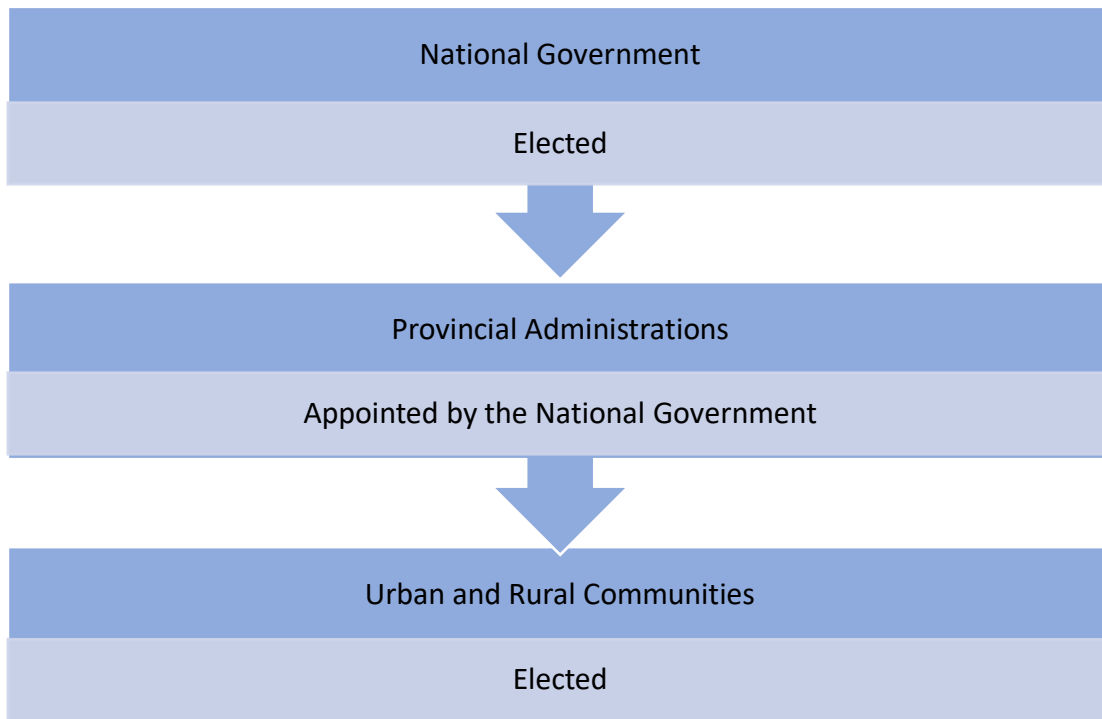
- Create a comprehensive framework of indicators for the results chains that would allow for measuring the important aspects of programs, such as effectiveness, efficiency, equity, cost effectiveness, access.
- Focus on the learning purpose of performance monitoring first and communicate it to the provincial administrations as an effort to understand how well programs work without ramification for the personnel. Once the indicator framework has been refined and proven to be useful for understanding performance, the Department could consider incorporating it in decision making related to compensation.
- Consider incorporating rigorous impact assessment in the evaluation system for programs that seem to be performing well based on their outcome data. This would be particularly useful for programs that the Department might want to scale up. If there are certain programs the Departments is interested to learn more about (for example, why they work, what are the important factors enabling success), they could consider qualitative evaluation.

Background information

The Republic of Armenia is a small country of 2.9 million people located in the South Caucasus region of the Eurasian continent. It declared independence from the Soviet Union in 1991.ⁱⁱ Armenia started off its new era of independence as a presidential republic, then transitioned into a semi-presidential system in 2005 and into a parliamentary system in 2015. Currently, the parliament elects the prime minister who forms the government and holds the executive power.

According to the Law on Administrative-territorial division adopted in 1995, the administrative-territorial units of Armenia are provinces and communities within provinces.ⁱⁱⁱ Armenia is divided into ten administrative regions (or provinces): Aragatsotn, Ararat, Armavir, Gegharkunik, Lori, Kotayk, Shirak, Syunik, Vayots Dzor and Tavush (see *Annex 2: Map of Armenia*). Heads of provincial administrations are appointed by the government; hence they are subordinate to the national government and do not exercise any autonomy. At the local level, the country has self-governing rural and urban communities. The largest community is capital city Yerevan (1.08 million), second largest city is Gyumri (113.1 thousand) and the third largest city is Vanadzor (78.1 thousand). Figure 1 illustrates the tiers of government in Armenia.

Figure 1. Tiers of government in Armenia



From 1997 through 2018, territorial administration was regulated by a Presidential Decree. Following the constitutional changes in 2015, this decree was declared void and the parliament adopted a law in November 2019 to regulate territorial administration in Armenia.^{iv}

According to the newly adopted law the principles of territorial administration are the following:

- harmonization of national and local interests
- cooperation among provincial administrations, other bodies of national government and their provincial branches, as well as local self-governing communities
- equitable and sustainable territorial development.

The Ministry of Territorial Administration and Infrastructure is responsible for developing the territorial policies of the government and the programs for implementing those policies. Provincial administrations are responsible for implementing the territorial policies of the government. The Ministry of Territorial Administration and Infrastructure is mandated to coordinate the work of provincial administrations in order to raise their effectiveness.

The law stipulates general and sectoral competences for provincial administrations. Being situated between the local self-governing municipalities and the national government an important part of their general competences is to facilitate the flow of information between the provincial communities and the national agencies (and their provincial branches) through collecting information relevant for territorial development, presenting provincial issues and

concerns to the government, making budgetary requests and suggestions on national programs related to provinces, etc.

In addition to general competences, the provincial administrations have responsibilities in the following sectors: a) finance, b) urban development, c) transport and road construction, d) agriculture and land use, e) education, f) health care, g) social protection, h) culture, sports and youth i) nature and environmental protection, j) defense h) civil protection and protection during emergency situations, i) economic development.^v Their sectoral responsibilities could be divided into the following big categories:

- implementation of government policies in the regions (for example, secondary education policies in educational institutions subject to the provinces, programs on protecting and improving the health of inhabitants, programs on preventing emergency situations or recovering from them, etc.)
- organizing the work of state institutions subject to the provinces (schools, health care, social protection, cultural institutions, etc.)
- construction and maintenance of public infrastructure (for example, provincial roads, buildings for educational, health care, social security, sports, cultural institutions, preparing the layouts of urban and rural communities, etc.)
- participation in national agencies' programs and activities in their provinces (for example, agricultural, environmental, anti-epidemic, defense)
- monitoring the situation in the abovementioned sectors
- preparation of the provincial development strategies and the annual activity plans, as well as monitoring of the implementation of those

- other responsibilities.

Based on the principles of territorial administration and the competences described above, programmatic documents are developed (development strategies, annual activity plans, etc.) that lay out the specific goals and results that provincial administrations are expected to achieve.

The existing performance evaluation framework of provincial administrations

Both the newly adopted law and the 1997 presidential decree include provisions on evaluating the work of provincial administrations. In 2009, the government introduced a performance evaluation system, the main principles and components of which were laid out in a corresponding government decision.^{vi}

The decision includes an annex of indicators, sub-indicators, grading instructions, and the national executive departments responsible for grading. Because provincial administrations have competences within the purview of many national executive departments, each department evaluates the part of the work that is in their domain.

The evaluation is performed annually. By February 15 of every year, the heads of provincial administrations submit reports to the Ministry of Territorial Administration and Infrastructure, which turns those over to the relevant national executive departments. The latter, together with the Council of Civil Service, examine those reports and submit their conclusions and grades to the Ministry of Territorial Administration and Infrastructure within seven days of receiving the reports. The grades are made public, they have to be posted on the official website of the RA government.

The current performance monitoring and evaluation system was introduced in 2009 according to the priorities and goals of the government that developed it. However, Armenia has gone through fundamental changes in the past few years. In particular, in 2018, a national peaceful mass movement removed president Serzh Sargsyan from office who attempted to maintain his grip on power by turning Armenia into a parliamentary country and being appointed the prime minister after serving two presidential terms.^{vii} A new parliament and government were formed following the revolution through the first free and fair elections in about twenty years. The current government has set out to reform the public administration system and strengthen Armenia's democratic institutions.^{viii} A part of that reform agenda is to make sure that the performance evaluation system of provincial administration reflects the goals of the current government. According to the Department of Territorial Administration at the Ministry of Territorial Administrations and Infrastructure those goals are as follows:

- to hold provincial administrations accountable to the government, as well as to the public by demonstrating the results of their work
- to incentivize better performance by measuring the results of their work

The Department is also open to learning how performance evaluation can help them learn about how programs work and to use that information for improvement.

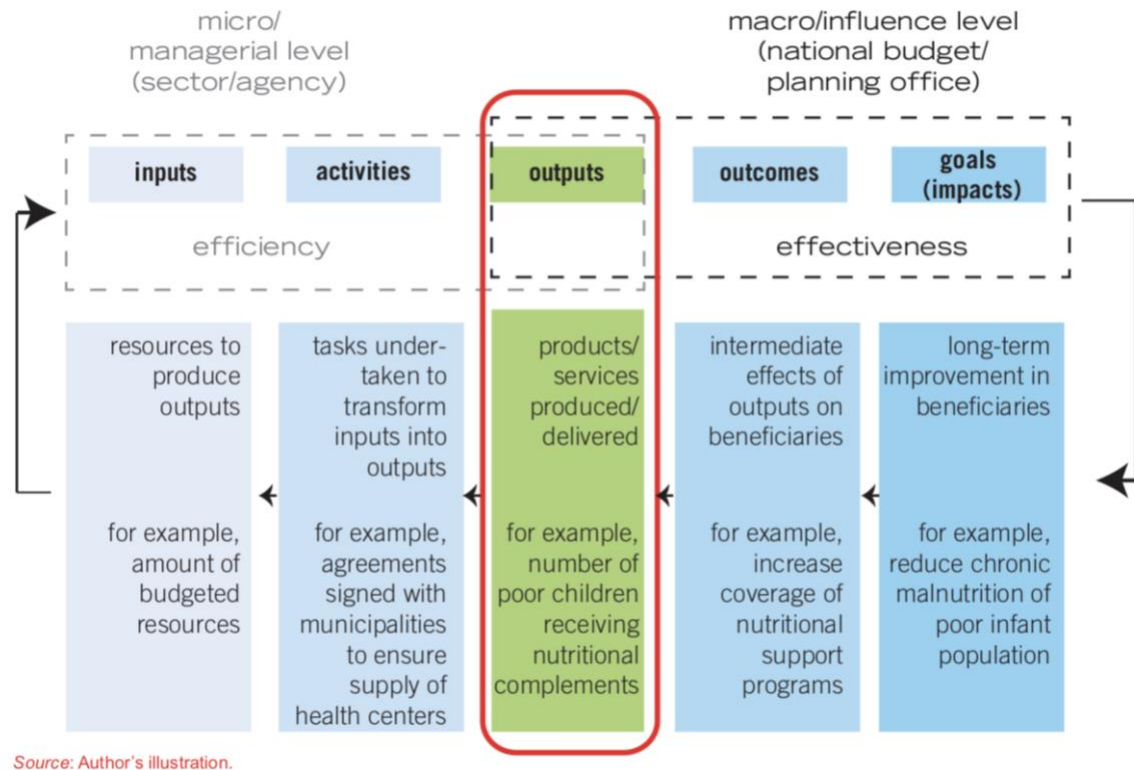
Research questions

There are different types of evaluation depending on the goal of evaluation. If the goal is accountability it would be necessary to measure different aspects of program results that various stakeholders care about (efficiency, effectiveness, equity, etc.). However, if policymakers are interested in learning about “why” and “how” programs work, measuring effect and efficiency would not be sufficient.^{ix} They would need to implement in depth qualitative evaluation to understand the reasons underlying performance.

Organizations and governments usually use log frames^x or logic models to monitor and evaluate the results achieved by their programs. While they can be modified to better suit the needs of a particular organization, most log frames follow a similar logic. For example, A World Bank Guide on Defining and Using Performance Indicators and Targets in Government Monitoring and Evaluation Systems states that it is necessary to establish a results chain for programs implemented by agencies and it should include identifying medium or long-run high level outcomes, as well as inputs, activities and outputs that would help achieve those outcomes. Figure 2 below illustrates a results chain:^{xi}

Figure 2: Performance Indicators and the Results Chain

Guide on Defining and Using Performance Indicators and Targets in Government Monitoring and Evaluation Systems, World Bank



The Guide also emphasizes the importance of setting targets for outputs and outcomes “to hold governments accountable and to exert external pressure over ministry or agency performance”.^{xii} It defines targets as “positive results-oriented statements that, starting from a baseline level, identify a path and destination (for example, increase coverage of nutritional support programs from 35 percent to 40 percent of poor people between 2008 and 2010)”. Finally, the Guide states that each target needs to have a performance indicator—“a quantitative or qualitative expression of a program or policy that offers a consistent way to measure progress toward the stated targets and goals”.^{xiii}

The goal of this analysis is to understand whether the existing evaluation system of Armenia's provincial administrations is suitable for promoting accountability and improving performance and how it can be modified to better serve the goals of the Department of Territorial Administration. In particular, I will look to see whether the existing performance evaluation system follows a logic similar to the results chain described above. In order to measure results, there have to be clearly defined results that provincial administrations are expected to deliver, and the evaluation system needs to have measures for all components of the result chain presented above to be able to measure effectiveness, efficiency, equity, etc. Therefore, this analysis seeks to answer the following questions:

- a. Do policies (relevant programmatic documents, such as multi-year development strategies, annual activity plans) specify clear goals, targets and outcomes that provincial administrations are expected to achieve? Are Inputs, activities and outputs tied to those outcomes in a results chain or logical framework?
- b. Which components of the results chain do evaluation indicators measure? Do they measure outcomes and impact?

Monitoring and evaluation emerged as widely used public administration tools in the 20th century by the pressures of economic, political and social forces. The expansion of welfare state in Europe and North America through 1960s led to implementation of ad hoc studies on social policies such as education, health care focused on analyzing their results and informing decision making. The economic crisis of 1970s pressed for decreasing public spending and evaluation oriented towards measuring efficiency. 1980s saw the rise of New Public Management, a wave of public administration reform aimed at incorporating private sector user-centered management practices into the public sector. These reforms also affected monitoring and evaluation practices. Governments started using indicators on inputs, activities, and outputs for accountability, and a greater emphasis was put on customer or user satisfaction along with efficiency. In mid 2000s the new trends on public administration, such as Public Value Governance, underscored the importance of democratic participation and inclusion, where citizens are not only viewed as beneficiaries, but active participants of policy making.^{xiv}

Developed countries and international organizations like the OECD, World Bank have made the most progress in terms of developing performance monitoring and evaluation systems that focus on holistic evaluation of policies (impact, equity, efficiency, cost-effectiveness, etc.), and not only on inputs or outputs. OECD defines monitoring as “A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.” Evaluation is defined

as “The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability.”^{xv}

The National Performance Management Advisory Commission (United States) in their Performance Management Framework for State and Local Governments writes that the transition towards performance management has been driven partly by the recognition that “A process approach to accountability is not sufficient. Officials, managers, and employees at all levels must be accountable not just for following processes but for producing results the public needs.” They emphasize that “Evaluation must be a component of performance management because understanding the relationship between the activities government carries out and the results it achieves is necessary to learning, improvement, and accountability.”^{xvi}

The New Zealand State Services Commission and The Treasury guide on Performance measurement states that “Performance measurement is a precursor to effective and informed management. It is crucial to agencies with policy, delivery, monitoring and/or sector oversight roles. It enables agencies and sectors to chart the progress they are making in improving outputs, outcomes and value-for-money, and to take corrective action if required.”^{xvii}

So, a common characteristic of monitoring and evaluation systems in many countries is that they have become more and more focused on outcomes or results in order to meet the growing demand for accountability and improvement.

In this section I will look into the programmatic documents that lay out the goals/expected results of the territorial development policies, the role of provincial administrations in achieving those goals, and will analyze the indicators used for their performance evaluation to see how they are connected with the results expected from them.

Do policies (relevant programmatic documents, such as multi-year development strategies, annual activity plans) specify clear goals, targets and outcomes that provincial administrations are expected to achieve? Are Inputs and activities tied to those outcomes in a results chain or logical framework?

The overarching document of Armenia's territorial development policy is the 2016-2025 Territorial Development Strategy^{xviii}, which analyzes the current situation, development challenges, the weaknesses and strengths of provinces and lays out long-run goals, priorities and targets. Based on this Strategy a 2018-2020 Territorial Development Operational Program^{xix} was developed, which specifies shorter-term targets and baselines, as well as lays out the framework for specific programs, projects and activities to be implemented in order to achieve those goals. Then, each province developed its own 2017-2025 Territorial Development Strategy^{xx} for achieving development goals, as well as annual activity plans for those strategies.

- [The 2016-2025 Territorial Development Strategy of the RA and the 2018-2020 Territorial Development Operational Program](#)

This document states the main goals and directions of Armenia's territorial development policy. The general goal of RA territorial development policy is the promotion of sustainable and equitable social and economic development in all provinces, taking into consideration the resources and needs of each province. This goal consists of the following components:^{xxi}

1. Increasing the competitiveness of all provinces based on their existing potential through creating opportunities for integration into national and international markets.
2. Guaranteeing maximum territorial equity, paying a special attention to the least developed and border areas, as well as to the sustainable use of resources.
3. Improving territorial development policies and practice, ensuring the active participation of territorial and local actors in planning and implementation.

The Strategy also includes the following targets against which the success of its implementation would be measured:^{xxii}

1. By 2025, the number of people with tertiary and higher education, employees in non-agricultural sectors, and the number of active enterprises will increase by 10% compared to 2014.
2. In 2025, per capita GDP in each province will exceed 60% of the average per capita GDP in Armenia and no more than 30% of provincial population will have less than 70% GDP per capita.
3. In 2017, the RA will have a general Territorial Development Operational Program for the implementation of territorial development policy, which will serve as a basis for designing budgetary programs and for receiving financial support from foreign sources.

- a. The first Operational Program will be developed for 2017-2020. By 2020, the first 40% of the targets will have to be achieved.
- b. The next Operational Program(s) will be developed based on the implementation of the previous program and will aim at archiving the remaining 60% of the targets.

The indicators of the first two targets are reasonable and clear measures for the corresponding first two goals, because they reflect the intended outcomes of policies. Measuring outcomes by those indicators would allow for assessing whether results have been achieved or not. The indicator of the third target is not tied to the corresponding goal as tightly, because the existence of a Territorial Development Operational Program will not necessarily lead to improved policies and practices, so its existence is not a measure of improvement. The indicator should measure the results that the Program intends to achieve. For example, if the purpose of developing that program is to design budgetary programs and attract financial resources than the indicator should be tied to those results. Moreover, there are no targets and indicators for the second part of the 3rd goal - ensuring the active participation of territorial and local actors in planning and implementation. While participation is a process-related goal, having targets and indicators would show government's commitment to it and it would be easier to demonstrate whether policies have been participatory.

The strategy also includes horizontal goals^{xxiii} (equality and non-discrimination, protection of competition, preservation of environment and energy sustainability, development of information society) without any evaluation indicators or targets, which makes it unclear what exactly is expected to achieve and how success will be measured. Taking into consideration that some of these might be competing goals, for example economic growth might take place at the expense

of preservation of environment and energy sustainability, it is especially important to have targets and indicators for these kind of goals in order to prioritize all of them.

The first Operational program mentioned under the third target of the Strategy was developed for 2018-2020. The goals laid out in this Program are similar to those of 2016-2025 Territorial Development Strategy and the Program also specifies the following targets and baselines, again in line with the targets of 2016-2025 Territorial Development Strategy:^{xxiv}

1. In 2020, the number of non-agricultural jobs in provinces will be at least 460,000, which is more than that in 2015 by 60,000 (15%)
2. In 2020, the number of active enterprises in the regions will be at least 40,000, which is more than that in 2015 by 1,200 (3%)
3. The proportion of people with tertiary and higher education will be 36%, which is a 5% increase from 2015.
4. The eight regions (all but Shirak and Tavush) exceed 60% of national GDP per capita maintaining the achievements in 2014 and 2015.
5. No more than 40% of provincial population will have less than 70% GDP per capita.
6. All territorial and local development projects for 2018-2020 will be consolidated under the 2018-2020 Territorial Development Operational Program. The latter will be entirely incorporated into the medium-term expenditure program.

- [2017-2025 Territorial Development Strategies of ten provinces](#)

Provinces have developed their own development strategies based on the 2016-2025 Territorial Development Strategy of the RA and the 2018-2020 Territorial Development

Operational Program. The Strategies of provinces also include goals mirroring the national level goals and targets to be met at the provincial level. Most outcomes are the same, but provinces often have more ambitious targets (for example, in Ararat province the target for the per capita GDP is 85% of the national average by 2025)^{xxv} and extra outcomes (for example, in Lori, the number of provincial administration employees with higher education).^{xxvi}

The Strategies also include the institutional framework for implementation. The stakeholders involved in the implementation are provincial administrations, local governments, the Ministry of Territorial Administration and Infrastructure, other ministries, donors, non-governmental organizations, the private sector. The responsibilities of provincial administrations with regard to implementing the Strategy are the following:

- Coordination of the holistic implementation of the Territorial Development Strategy
- Development, revision and adjustment of the Territorial Development Strategy as needed
- Coordination of activities related to the implementation of Territorial Development Strategy
- Preparation of annual reports on the implementation of Territorial Development Strategy
- Providing all stakeholders with information on the implementation of Territorial Development Strategy and other issues related to provincial development
- Providing municipalities with professional and technical support on developing their own development strategies, as well as project proposals
- Supporting public unions and other stakeholders in preparing local and regional development projects
- Promoting inter-city cooperation aimed at local development and project planning

- Promoting the development potential of the province

So, while other stakeholders are responsible for certain components of the strategy, provincial administrations have more overarching responsibilities of developing the strategies, coordination of its implementation and supporting other stakeholders. This multi-stakeholder involvement in the implementation of strategies can make it challenging to create indicators for the provincial administrations' activities. British Columbia's Office of the Auditor General touches upon this issue in their Guide for Developing Relevant Key Performance Indicators (KPIs) for Public Sector Reporting. The Guide says "Some government organizations do not provide services directly. Rather, they guide and fund agencies that deliver the services on government's behalf. This can create a challenge for those organizations when it comes to identifying relevant KPIs that reflect the reporting organization's activities. However, the type of outcomes an organization wants to achieve from its activities is still the same even when the organization is not directly providing the service or program. Only the degree of influence, tools and activities used to achieve the goal is different. Therefore, organizations in these situations should still create KPIs to evaluate their performance towards the outcomes they seek to achieve."^{xxvii}

- [Annual Activity Plans for the 2017-2025 Territorial Development Strategies of provinces](#)

Provinces develop annual activity plans for the implementation of 2017-2025 Territorial Development Strategies.^{xxviii} The activity plans include strategic goals mostly aligned with the goals in development strategies, brief descriptions of activities/programs aimed at achieving those goals, indicators/targets, sources of information, responsible bodies, time periods, as well as risks and assumptions. For some programs the indicators measure outputs and outcomes. For

example, in Ararat region the program for providing assistance to existing enterprises in the 2018 annual activity plan has the following indicators:^{xxix}

Outputs

- Number of enterprises participating in needs assessment
- Number of trainings
- Number of training participants
- Number of consulting projects
- Number of enterprises participating in consulting projects
- Number of networking events
- Number networking event participants
- Number of export agreements

Outcomes

- Growth of exports
- Number of newly created jobs

While these indicators would help measure the effectiveness of activities, they are not enough to measure access to these services (for example, percent of enterprise in need participating in the programs), equity (for example, the urban rural breakdown of enterprises), efficiency (for example, cost per unit of output), or cost-effectives (for example, cost per unit of outcome). Moreover, there is no separation of output and outcome indicators, they are all in one column (the separation above is that of the author). So, it would be useful to have separate columns for outputs and outcomes, as well as to add inputs and relevant indicators to measure equity, access, etc.

Some activities only have output indicators. For, example, in the 2018 annual activity plan of Syunik region^{xxx}, the indicator for a teacher quality improvement program is the percent of teachers participating in trainings and there is no outcome indicator measuring the extent of improvement.

There are some indicators for strategic goals and/or programs that are activities. Some indicators do not seem to reflect the goals they are meant to measure. For example, the 2018 annual activity plan of Shirak province has the following strategic goal: “Increasing the involvement of provincial and local actors in the implementation of provincial development policy and improving human resource management skills”.^{xxxi} The indicator for this goal is: “Increasing residents’ satisfaction with the quality of services provided by municipalities”. This is worded as an activity, and it is unclear how satisfaction will be measured. Moreover, it is arguable whether residents’ satisfaction with the quality of services is an adequate measure of the inclusiveness of policy implementation.

To sum, while some goals and activities in the annual activity plans have output and outcome indicators, some of them don’t have any (or have inadequate indicators) and some only have output indicators. There is no comprehensive framework of indicators that would allow for measuring efficiency, effectiveness, equity and other important aspects of programs.

- [Annul activity plans of provinces](#)

The day to day activities of provincial administrations are guided by their annul activity plans, which include activities that would presumably help achieve the targets stated in the above-mentioned documents, along with activities not directly related to those strategies. These

activity plans include goals, indicators and outputs, but no targets and outcomes. Below are two examples.

Table 1. Excerpts from 2018 Annual Activity Plans

Activity	Goal	Indicator	Output
Assistance to farmers ^{xxxii}	Providing farmers with the necessary amount of seeds, fertilizers, fuel, pesticides, and spare parts	According to established procedures submitting requests to the relevant national executive bodies for high quality seeds, fertilizers of necessary quantity and quality, fuel, pesticides, and spare parts	Relevant documents
Oversight over the implementation of national programs on social security ^{xxxiii}	Improvement of the conditions of low income and other vulnerable people	Collecting data on vulnerable groups and presenting those to the relevant national executive bodies, existence of an updated databases by each vulnerable group	Relevant documents

Goals are often activities as in the first example in the table above. Moreover, it aims at providing farmers with materials rather than improving their outcomes, such as harvest or profits. Providing farmers with those materials would presumably improve their outcomes, but if the goal is not tied to the outcome itself, the provincial administrations would be able to demonstrate progress towards the implementation of the activity regardless of its impact on the outcome, when the outcome is what matters for stakeholders.

Indicators often do not directly reflect the goals as illustrated in the second example above, where the goal is to improve the conditions of vulnerable groups, but the indicator is collection of data on them. In this example, while information is very important, the existence of data is not a metric of improvement.

Another key shortcoming of the annual activity plans is that they do not incorporate the targets and outcomes specified in the strategic development documents described above. This disconnect means that the success of the day to day activities or provincial administrations is not guided and measured by the outcomes they are expected to achieve.

Key conclusions

- Strategic development documents state goals, outcomes and targets to be achieved, although there are horizontal goals, which do not have targets and outcomes.
- There are multiple stakeholders involved in the implementation of those strategies. The provincial administrations have a central in it, including developing the provincial development strategies, coordinating their implementation, and assisting other actors.
- The annual activity plans of strategic documents do not have a comprehensive framework of indicators that would allow for measuring efficiency, effectiveness, equity and other important aspects of programs.
- The annual activity plans of provincial administrations lack outcomes, outcome indicators and targets. Most indicators track inputs or outputs, so there is a disconnect between the activities of provincial administrations and the outcomes expected from them.

Which components of the results chain do evaluation indicators measure? Do they measure outputs, outcomes and impact?

The evaluation system that is currently used includes 18 main indicators in total. Fifteen of them are phrased as “the completeness of provincial administration’s work” in the fields, where

they have competences (health care, education, agriculture, etc.). The remaining three of the main indicators are: “the application of civil service legislation in provincial administrations”, “cooperation with international organizations and bilateral relations”, “actions taken for the maintenance and effective use of state property”. Each main indicator has several first-level sub-indicators. While the government decision calls them indicators, neither the main indicators, nor most sub-indicators are indicators, because most of them are worded as activities. For example, “actions taken for the maintenance and effective use of state property” implies that what it measures are the “actions” and not how well state property was maintained and used.

National executive departments responsible for evaluating the work of provincial administrations have developed guidelines^{xxxiv} that include second-level sub-indicators for the abovementioned indicators and instructions for grading. While there are some second-level sub-indicators that would measure other important aspects of programs (such as, compliance), most of them follow the pattern of reflecting activities. Below are some examples to illustrate.

Finance

The Ministry of Finance guidelines outline 8 second level sub-indicators for 3 out of 5 first level sub indicators. Most of the second level sub indicators are technical activities (for example, receiving requests for subventions from the communities and providing those to the RA Ministry of Finance) unclear to what end. The indicators for procurement are worded as activities, but they are in essence compliance indicators (for example, ensuring that procured products correspond to the product descriptions approved by the provincial administration).

Health Care

The guideline developed by the Ministry of Health Care specifies 17 second level sub-indicators for the 4 first level sub indicators. Most second level sub indicators are activities that provincial administrations have to implement. For example, participation in the streamlining projects of health care system; supporting the implementation of sanitary-hygienic, anti-epidemic and quarantine activities; maintenance and technical equipment of the buildings of provincial health care institutions etc. Two of the very few sub indicators that could serve as such are the following: 1. absence of debt of salaries and social payments, 2. the trends of disease and death indicators in the province in the past three years.

Some of the activities include no outcomes (for example, organizing the work of provincial health care organizations), and it is unclear what they are meant to achieve. Others have outcomes (for example, taking actions aimed at improving the quality of services in primary health care institutions in the province), however the indicators are not meant to measure the outcomes themselves (improved health care services, in this example), but the actions taken by the provincial administrations.

Social Security

For three of the four first level sub indicators the guidelines lay out subsectors as second level indicators, such as social assistance, elderly and persons with disability, etc. These broad subsectors would be very hard to evaluate. The fourth first level sub indicator includes activities under subsectors. Some of the activities do not include a clear outcome (for example, coordination of the provincial committees for issues related to persons with

disability). For some, it is clear what the outcome is (for example, implementation of activities of the program on ensuring equal opportunities for men and women), but it does not serve as the measure of success.

The fact that the system is based on activities rather than qualitative or quantitative indicators is a serious shortcoming. It would not create incentives for provincial administrations to produce better outcomes and be accountable, because implementation of activities would not necessarily lead to improvement of services or achievement of targets. Moreover, evaluating activities would involve considerable subjective judgement and would not be reliable.

The 2009 Government decision on performance evaluation established a grading system, which states maximum points for each indicator and sub-indicator. The points for indicators sum up to 100. Up to 60 points was considered unsatisfactory, from 60 to 85 points would be satisfactory, and above 85 points would be considered good performance. In 2010, all three thresholds were raised presumably to set higher expectations and promote better performance. The new thresholds were 75 and 90 – below 75 would be unsatisfactory, from 75 to 90 would be satisfactory and above 90 would be good.

In 2009, the lowest grade was 81.2, the average grade was 87.3 and the highest grade was 93.5, eight out of ten provincial administrations received good grades, and the remaining two received satisfactory grades. After the change in threshold, in 2010, the lowest grade was 91.6, the average grade was 93.4 and the highest grade was 96.08.^{xxxv} This time the performance of all provinces was assessed to be good. Since then the average grade has always been above 90 and the grades of most provincial administrations have also been above the threshold for good. The fact that after the change in threshold, all provincial administrations received higher grades than

the year before suggests that grading was quite subjective. In fact, the guidelines for grading are often broad and leave a lot of room for subjective judgement. Below are some of the guidelines to illustrate.

Finance

The grading for most indicators is based on timeliness, quantity and quality. Here quality is defined as compliance with established procedures. While compliance with procedures is an important criterion to keep track of, it would not necessarily measure quality by itself. The grading for procurement second-level sub indicators is based on the percent of procurement procedures and reports approved by the relevant national agency, which makes it appear more credible.

Health Care

The grading is based on quality of implementation and timeliness.

The criteria of quality are the following:

- Implemented without flaws
- Implemented with non-essential flaws
- Implemented with essential flaws
- Not implemented

The criteria of timeliness are the following:

- Timely implementation
- Non-essential violation of deadlines
- Essential violation of deadlines

The guidelines do not specify how the quality of implementation will be assessed, in particular, what constitutes flawless implementation, how flaws will be classified as essential or non-essential. Similarly, it does not specify how violations of deadlines would be classified as essential or non-essential.

Land Use

Grading guidelines require to scale the scores for each sub indicator by a certain percent of activities implemented (for example, if up to 30% of communities have made targeted use of the land resources, they would be given 0 points, if more than 30% of communities have made targeted use of the land resources, they would be given full points). However, there is still space for subjective interpretation. In this particular example, it is unclear how they would determine whether the use of resources was targeted or not. Moreover, the threshold for receiving full points seems to be very low.

Protection of the Environment

The guidelines do not include any instructions for grading, but they require provincial administrations to submit information on extra actions taken to better implement the activities, description of activities that were not implemented or were not properly implemented, including explanations. These requirements might incentivize the provincial administrations to perform better.

Culture

The grading guidelines list several activities for each second-level sub indicator and instruct to award full points if all activities are implemented and partial points if some of

them are implemented. For example, for the promotion of modern art and culture, the following activities are listed:

- Assist new performances and concerts based on classical and contemporary playwriters and composers
- Promote the accessibility/affordability of exposure to folk culture and art
- Create equal opportunities for searching new ideas and values, the use of the societies' creative potential

If two of these activities were implemented, the provincial administration would be awarded 0.15 out of 0.3 points. If less than two activities were implemented, they would not receive any points for this sub indicator.

There is no mentioning of quality or timeliness in the grading guidelines, hence they evaluate solely based on the number of activities implemented. Moreover, there are no further guidelines on how the ministry would determine whether an activity was implemented or not leaving it to the discretion of the individual evaluator.

Key conclusions

- The monitoring and evaluation system lacks indicators to monitor and evaluate outputs and outcomes, and mostly evaluates activities. The targets and outcomes laid out in the programmatic documents are not reflected in the evaluation system.
- The guidelines for evaluation of the implementation of activities tend to be vague and they leave plenty of room for discretion.

Examples of performance monitoring and evaluation systems

The experience of other countries that have made considerable progress in developing performance monitoring and evaluation systems can be useful. Below are three examples from the United States and Australia, which are quite comprehensive and robust systems.

North Carolina Benchmarking Project, USA

The North Carolina Benchmarking Project ^{xxxvi} was initiated in 1995 by a group of municipalities, the University of North Carolina School of Government and North Carolina Local Government Budget Association. After three phases of piloting, the project has been implemented on an annual basis since 2000. The project publishes reports^{xxxvii} on key service areas of participating municipalities (Asphalt Maintenance & Repair, Police Services, Residential Refuse Collection, etc.). The reports comprise of standardized, two-page information on each service area and municipality. They provide information on the following:

1. “Performance Measures. Three types of performance measures are used and reported: workload, efficiency, and effectiveness.
 - a. Measures of Workload: These measures address the level of service being provided or provide an assessment of need for a service. An example of a workload measure is tons of residential refuse collected.
 - b. Measures of Efficiency: These measures may include cost per unit of service provided, cost per unit of output, or the units of service provided per full time

equivalent position. An example of an efficiency measure is cost per ton of residential refuse collected.

- c. Measures of Effectiveness: These measures refer to the quality of service provided, citizen perceptions of quality, or the extent a service meets the need for which it was created. An example of an effectiveness measure is complaints per 1,000 collection points of residential refuse
2. Municipal Profile. This includes a limited number of characteristics of each jurisdiction, such as population density and unemployment rate, which may affect service performance and cost.
 3. Full Cost Profile. A cost accounting model is used to calculate full or total cost of providing each service area under study. Although the cost data were collected in detail, using a collection instrument with more than seventy specific line items, the reporting format aggregates the detailed cost data into three general categories for the purpose of presentation: personal services for the direct expenses of salaries, wages, and related fringe benefits; operating costs that include direct operating expenses and indirect cost allocations; and capital costs that represent usage charges for equipment and facilities.
 4. Service Profile. This identifies important dimensions of service organization and method of delivery.
 5. Explanatory Information. This segment of the report describes how the service is delivered and identifies conditions or dimensions that affect performance and cost data of the service." xxxviii

Participating municipalities have used the data for service improvement, for example in residential refuse collection and household recycling. One of the key learnings of the project is that “Auditing or verifying the accuracy of performance data is a necessary component of performance measurement and benchmarking”.^{xxxix}

County of San Mateo, California, USA

The county has created online dashboards that allow citizens to access progress reports, as well as trends, raw data and actions taken to improve performance.^{xl} Performance information is presented for the programs and services provided by each department of the county. The dashboard for a service area or program typically includes a brief overview of the service, program outcome statement, program services (in bullet points), performance data, the story behind performance and future priorities.

For example, the program outcome for Public Works: Road Construction & Operations^{xli} is “Ensure that County maintained roads are safe, accessible and well maintained”. The indicators are measures of quantity, efficiency and quality of services:

- Volume of Trash Collected per Mile Swept
- Cost per Mile of Road Maintenance
- Percent of Maintained Miles with Pavement Condition Indexes (PCI) Greater than Baseline

The program outcome for San Mateo Medical Center (SMMC)^{xlii} is the following: “We partner with patients to provide excellent care with compassion and respect.” The indicators for this service are measure of patient satisfaction, efficiency and access to health care services:

- Likelihood to Recommend SMMC to Family and Friends
- Cost Per Patient Per Month
- Number of Patients Assigned to SMMC by Health Plan of San Mateo (HPSM) Who Have Not Been Seen

Performance data is presented with graphs that show the trends in the past five-six years and the targets intended to achieve. These dashboards give the viewer a brief, but holistic view of the services provided, key aspects of performance, explanations on performance and planned actions to address underperformance or improve services.

It is important to note that the County tracks data on five community outcomes and nine goals for its Shared Vision 2025 Strategy that are broader, such as unemployment rate, poverty, greenhouse gas emissions, recidivism rates, etc.^{xliii} The dashboards provide data on the general situation in a particular area in the county, as well as information on the services provided by the county that could contribute to the trends. For example, for unemployment, the county tracks Annual Average Unemployment Rates, Unemployment Rates by City and Census Designated Place, etc., and they provide brief information on their services aimed at decreasing unemployment, including a graph on the Percent of Employment Services Clients Placed in Unsubsidized Employment.^{xliiv} While the general trends cannot be entirely or easily attributed to the work of the county, the goals of the county are to improve broader community outcomes, therefore it is important for it to track those outcomes and to tie their services and program indicators to those broad community outcomes.

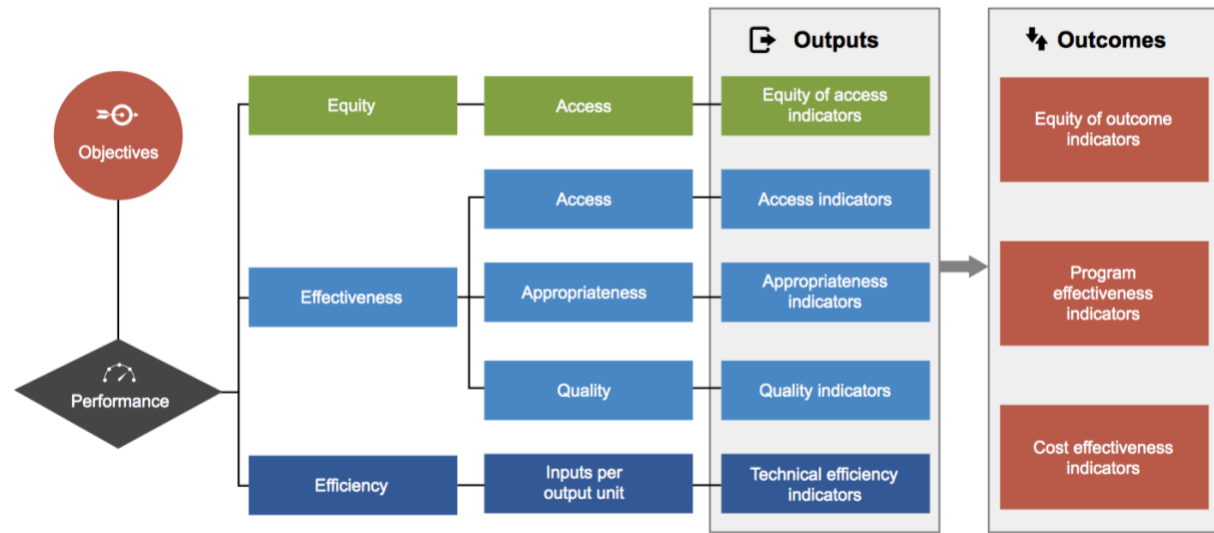
The annual Report on Government Services (RoGS) provides information on the equity, effectiveness and efficiency of government services in Australia.

Australia has a quite comprehensive performance monitoring and benchmarking system. The inter-governmental Steering Committee for the Review of Government Service Provision is responsible for reporting on government performance with the support of the Productivity Commission, which is the Australian Government's independent research and advisory organization.^{xlvi} The annual reports on government services are used to inform evaluation and decision making (planning, budgeting), as well as to demonstrate accountability. The report provides information across jurisdictions, which allows for comparison of government performance across states and territories. It focuses on social services, such as childcare, education and training, health, justice, emergency management.

Below is the general Framework of performance indicators, which is adapted for each service area to better serve its unique priorities and needs.

Figure 3: General Performance Indicator Framework

Australian government, Productivity commission: Report on Government Services



The government sets objectives against which performance is measured. The objectives have the following components:

- high-level goals that reflect the intended impact on the community
- objectives related to characteristics of effective service delivery
- objectives related to equitable and efficient service delivery

The efficiency, effectiveness and equity of service delivery are measured by output indicators. Jurisdictions have similar objectives, however different objectives can be prioritized in different jurisdictions, for example one state might focus on improving accessibility, while another state might prioritize improving quality.

Equity of access indicators compare the proportion of people belonging to special groups (people living in rural and remote areas, non-English speaking people, etc.) in a community to their proportion in the pool of service beneficiaries in that community. For needs-based services, comparisons are made across needs-groups.

Access indicators measure overall access (how readily services are accessed by those in need relative to the relevant population, for example number of people with disability who accessed a specialist service compared to the number of people with disability), timeliness of access (for example, waiting times), affordability (for example, proportion of income spent on a service).

Appropriateness indicators measure the extent to which a service is appropriate for the needs of beneficiaries who have access to it and how well the service meets those needs (for example, whether students participating in a training achieve their goal).

Quality indicators measure whether services meet required standards (for example, compliance with legislation), their safety (for example, road safety), responsiveness (for example, patient satisfaction) and continuity (for example, whether services were uninterrupted).

Technical efficiency measures the cost of producing goods and services (output vs input).

High-level objectives have outcome indicators that provide information on the equity, effectiveness and cost-effectiveness of the impact of programs on the society in the short, intermediate or long run. The authors of the report acknowledge that outcomes are often difficult to measure, and they are more likely to be affected by external factors beyond the control of service providers. Where possible, they point to other factors and data to explain certain outcomes.

Recommendations

While strategic documents set development goals and outcomes, there is a disconnect between the activities of provincial administrations and the end-results expected from them. The key shortcoming of the performance monitoring and evaluation system of provincial administrations is that it is based on activities, the implementation of which is evaluated according to vague guidelines; therefore, the results have high degree of subjectivity and cannot meaningful contribute neither to accountability nor to learning and improvement. In order to address this key shortcoming, it will be necessary to move away from an activity-based system to output and outcome-based annual activity plans and a performance evaluation system. It is important that the new metrics are set through an inclusive process, in consultation with all relevant stakeholders (provincial administrations staff, private sector, citizens, civil society, etc.) to get a buy-in from all parties. Below are the suggested steps the Department of Territorial Administration could take moving forward.

- Create result chains for the programs outlined in annual activity plans of strategic development documents, where goals, inputs, activities, outputs and outcomes would be logically connected to each other. Make sure that outputs and outcomes are not worded as activities and are in fact output and outcome statements.
 - It is often difficult to establish a causal relationship between activities implemented and outcomes for the society, because high-level outcomes can be affected by external factors beyond the control of implementors. However, it is important to link the activities to societal outcomes, because the ultimate

purpose of the public sector is to improve outcomes for the society. At any given moment it will be difficult to attribute outcomes to activities without a randomized control trial but monitoring trends over time would allow for seeing correlations between outputs and outcomes. Correlations would not indicate causation, but they would still help make better informed decisions than in the absence of any data.

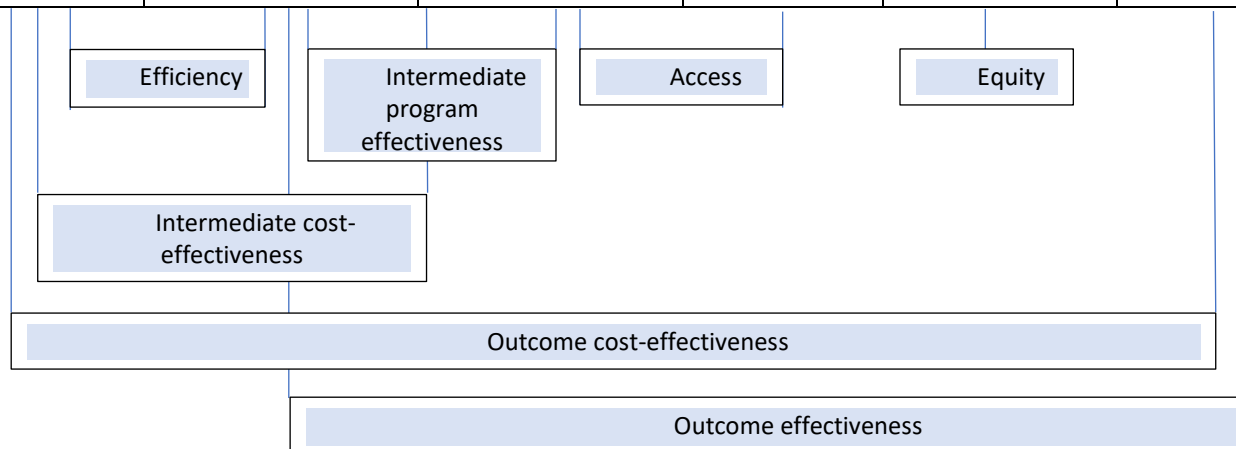
- Intermediate outcomes might also be affected by external factors to some extent, but they should be more easily attributable to the work of provincial administrations, for example, in economic development, an intermediate outcome could be increased investments in provinces, and the indicator would be the amount of investments attracted by provincial administrations.
- Create a comprehensive framework of indicators for the results chains that would allow for measuring the important aspects of programs, such as effectiveness, efficiency, equity, cost effectiveness, access. Annex 1 is a suggested framework for a province, which could be used for all strategic goals and programs aimed at achieving those goals. It incorporates indicators that would allow for tracking inputs, outputs, intermediate and high level/long run outcomes, and other relevant components. These metrics would then allow for measuring efficiency, intermediate program effectiveness, intermediate cost-effectiveness, access and equity. Final outcome effectiveness and cost-effectiveness would typically be possible to measure through randomized control trials or quasi-experimental techniques.

- There might be multiple possible indicators for any component of the framework (inputs, outputs, outcomes, etc). Since collecting data is costly and time-consuming, it is better to concentrate on a few or a single most important indicator.
- Data for some of the indicators might not be readily available. For example, if teacher trainings do not implement pre and post assessment, there will be no data to measure increase in teacher quality. For these cases, it will be necessary to incorporate extra data collection procedures in activities. While it will be tempting to use indicators for which data is already available, it is important that the metric measures what it is meant to measure, for example output indicators should not be used to measure outcome.
- Evaluation could be tied to compensation and promotion. Intermediate outcomes should be easier to use for that purpose than long-term or high-level outcomes, because the latter are much more susceptible to external influences. However, it is important to recognize the tradeoff between learning and accountability. If evaluation results are tied to salaries it might incentivize fraud in reporting, especially if there are no third-party data validation mechanisms. Therefore, it might be better to focus on the learning purpose of performance monitoring first and communicate it to the provincial administrations as an effort to understand how well programs work without ramification for the personnel. Once the indicator framework has been refined and proven to be useful for understanding performance, the Department could consider incorporating it in decision making related to compensation.

- The existing grading system could be abandoned, because it does not have the necessary rigor and objectivity to serve its purpose. In fact, comparing indicators across provinces would allow for understanding how well each province is doing compared to others, as it is in the case of the North Carolina Benchmarking Project described above. In that sense there is no need for grading. One reason, why awarding grades to provincial administrations might be helpful, is to use those grades for compensation-related decisions. In case, the Department wishes to do this in the future, it should be straightforward to use intermediate outcome indicators to construct grades weighted by achievement measured by relevant indicators.
- Along with comparing the data across provinces, it would be useful to compare those over time to follow yearly (or more frequent - quarterly) trends.
- The data could be summarized in annual reports. Countries also create online dashboards to present their performance data with graphs and chart, as well as they make the raw data accessible to the public, which is a good practice to follow suit.
- Moving forward, the Department could consider incorporating rigorous impact assessment in the evaluation system for programs that seem to be performing well based on their outcome data. This would be particularly useful for programs that the Department might want to scale up. If there are certain programs the Departments is interested to learn more about (for example, why they work, what are the important factors enabling success), they could consider qualitative evaluation.

Annex 1: Suggested Indicator Framework for a province

Strategic goal	Program	Input	Output	Intermediate outcome	Pool of potential beneficiaries	Rural/urban breakdown	Outcome
Promoting economic growth in the province	Attracting investments	Total cost	Number of investment packages presented to investors	Amount of investments attracted		% of rural investments	Per capita GDP of the province as a percent of average per capita GDP in Armenia
	Improving skills of unemployed people	Total cost	Number of people participating in trainings	Number of trainees employed within 3 months	Total number of unemployed people	% of rural trainees employed within 3 months	
Improving the quality of educational services in the province	Improving teacher quality	Total cost	Number of teachers participating in training	Average increase in teacher quality	Total number of teachers in the province	Average increase in rural teacher quality	Average increase in student achievement measured by standardized tests
	Increasing access to pre-school education	Total cost		Number of children newly enrolled in pre-school	Total number of pre-school age children	% children in rural communities	



Annex 2: Map of Armenia



Endnotes

ⁱ Making Measures Meaningful, The Logical Framework Approach, PM² Connections Performance measurement & management, Government Finance Review, October 2013,

https://www.gfoa.org/sites/default/files/GFR_OCT_13_63.pdf

ⁱⁱ About Armenia, Official web site of the Government of Armenia, <https://www.gov.am/en/official/>

ⁱⁱⁱ RA Law on Administrative-territorial division, 1995 <https://www.arlis.am/DocumentView.aspx?docid=137947> (in Armenian)

^{iv} RA Law on Territorial Administration, 2019, <http://parliament.am/legislation.php?sel=show&ID=6822&lang=arm> (in Armenian)

^v Decree of RA President on government in the regions of the Republic of Armenia, <https://www.arlis.am/DocumentView.aspx?docid=51838> (in Armenian)

^{vi} RA Law on Evaluating the Performance of Provincial Administrations, Performance Indicators, as well as the Grades Based on those, N 562-N, 2009, <https://www.arlis.am/DocumentView.aspx?docID=95807> (in Armenian)

^{vii} Armenia's Democratic Dreams, Anna Ohanyan, Foreign Policy, November 2018, <https://foreignpolicy.com/2018/11/07/armenias-democratic-dreams/>

^{viii} Armenia election: PM Nikol Pashinyan wins by landslide, BBC, December 2018, <https://www.bbc.com/news/world-europe-46502681>

^{ix} Performance Management Framework for State and Local Governments, The National Performance Management Advisory Commission, page 38,

<https://www.gfoa.org/sites/default/files/APerformanceManagementFramework.pdf>

^x The logframe as a monitoring and evaluation tool for government interventions in a chaotic and complex environment, Dominique E. Uwizeyimana, 30 Jan. 2020, <https://apsdpr.org/index.php/apsdpr/article/view/328/524>

^{xi} Defining and Using Performance Indicators and Targets in Government M&E Systems, World Bank, page 2 <http://documents.worldbank.org/curated/en/518971468346443826/pdf/643870BRI0Defi00Box0361535B0PUBLICO.pdf>

^{xii} See xi

^{xiii} See xi

^{xiv} "Comparative study on the institutionalization of evaluation in Europe and Latin America", Blanca Lázaro, Study n. 15, Series: State of the Art Area: Public Finance, July 2015, Programme For Social Cohesion in Latin America funded by the European Union, pages 21-26, http://sia.eurosocial-ii.eu/files/docs/1456851768-E_15_ENfin.pdf

^{xv} Glossary of Key Terms in Evaluation and Results Based Management, OECD 2010, <https://www.oecd.org/dac/evaluation/2754804.pdf>

^{xvi} Performance Management Framework for State and Local Governments, The National Performance Management Advisory Commission, page 4,

<https://www.gfoa.org/sites/default/files/APerformanceManagementFramework.pdf>

^{xvii} Performance Measurement: Advice and examples on how to develop effective frameworks, New Zealand State Services Commission and The Treasury, page 7, <https://ssc.govt.nz/assets/Legacy/resources/performance-measurement.pdf>

^{xviii} The RA 2016-2025 Territorial Development Strategy, <http://www.mtad.am/files/docs/1770.pdf> (in Armenian)

^{xix} The 2018-2020 Territorial Development Operational Program <http://www.mtad.am/files/docs/2049.pdf> (in Armenian)

^{xx} The 2017-2025 Territorial Development Strategies of provinces, <http://mtad.am/hy/marzerirazmavarutyun/>

^{xxi} See xviii, pages 59-60

^{xxii} See xviii, page 65

^{xxiii} See xviii, page 62

^{xxiv} See xviii, pages 69-70

^{xxv} The 2017-2025 Territorial Development Strategies of Ararat province, pages 48-49, <http://www.mtad.am/files/docs/1772.pdf> (in Armenian)

-
- ^{xxvi} The 2017-2025 Territorial Development Strategies of Lori province, pages 45-46, <http://www.mtad.am/files/docs/1772.pdf> (in Armenian)
- ^{xxvii} Guide for Developing Relevant Key Performance Indicators for Public Sector Reporting, Office of the Auditor General, British Columbia, Canada, page 7, https://www.bcauditor.com/sites/default/files/publications/2010/report_10/report/OAGBC_KPI_2010_updated.pdf
- ^{xxviii} 2018 Annual Activity Plans for the 2017-2025 Territorial Development Strategies of provinces, <http://mtad.am/hy/tarekancragir/> (in Armenian)
- ^{xxix} 1st Appendix of the 2018 Annual Activity Plan for the 2017-2025 Territorial Development Strategy of Ararat province, <http://ararat.mtad.am/development-plan/> (in Armenian)
- ^{xxx} Appendix of the 2018 Annual Activity Plan for the 2017-2025 Territorial Development Strategy of Syunik province, <http://syunik.mtad.am/development-plan/> (in Armenian)
- ^{xxxi} See xxx
- ^{xxxii} 2018 Annual Activity Plan of Aragatsotn Province, <http://aragatsotn.mtad.am/MATs/> (in Armenian)
- ^{xxxiii} 2018 Annual Activity Plan of Gegharkunik Province, <http://gegharkunik.mtad.am/work-plan/> (in Armenian)
- ^{xxxiv} The guidelines have been accessed on Armenia's Legal Information Platform (official website <https://www.arlis.am>) or requested from national executive bodies.
- ^{xxxv} The evaluations are available for years 2009 through 2013 and 2017. <http://mtad.am/hy/gnahatum/>
- ^{xxxvi} North Carolina Benchmarking Project, History, <https://www.sog.unc.edu/resources/microsites/north-carolina-benchmarking-project/history>
- ^{xxxvii} Report for the Fiscal year 2008-2009, <https://durhamnc.gov/DocumentCenter/View/539/North-Carolina-Benchmarking-Project-PDF?bidId=>
- ^{xxxviii} North Carolina Benchmarking Project, Measures, <https://www.sog.unc.edu/resources/microsites/north-carolina-benchmarking-project/measures>
- ^{xxxix} North Carolina Benchmarking Project, Achievements and Learning <https://www.sog.unc.edu/resources/microsites/north-carolina-benchmarking-project/achievements-and-learning>
- ^{xl} County of San Mateo, Performance, <https://performance.smcgov.org>
- ^{xli} County of San Mateo, Public Works: Road Construction & Operations, <https://performance.smcgov.org/stories/s/Public-Works-Road-Construction-Operations-4520B-/6sju-gm3q>
- ^{xlii} County of San Mateo , Health System: San Mateo Medical Center, <https://performance.smcgov.org/stories/s/San-Mateo-Medical-Center-6600B-/htng-edcu>
- ^{xliii} County of San Mateo , Shared Vision 2025 Dashboard, <https://performance.smcgov.org/shared-vision>
- ^{xliv} County of San Mateo, Unemployment Data <https://performance.smcgov.org/en/stat/goals/single/vpf3-n683/view>
- ^{xlv} Report on Government Services 2019, Approach to performance measurement, https://www.pc.gov.au/research/ongoing/report-on-government-services?id=141009&queries_year_query=2019&search_page_191702_submit_button=Submit¤t_result_page=1&results_per_page=0&submitted_search_category=&mode=results
- ^{xlvi} Australian Government Productivity Commission, Core Functions, <https://www.pc.gov.au/about/core-functions/performance-monitoring>