



Organisation for Economic Co-operation and Development

GOV/PGC/SBO(2020)4

Unclassified

English - Or. English

30 March 2020

**PUBLIC GOVERNANCE DIRECTORATE
PUBLIC GOVERNANCE COMMITTEE**

Cancels & replaces the same document of 26 March 2020

Working Party of Senior Budget Officials

Chile: Review of DIPRES Programme Evaluation System

JT03460195

Acronyms

BIPS	Integrated Bank of Social Programs
CBA	Central Budget Authority
CGR	Comptroller General of the Republic of Chile
DIPRES	Budget Office of Chile
EFA	Focused Scope Evaluation
EIG	Institutional Expenditure Evaluation
EPG	Evaluation of Public Programs
EPN	Evaluation of New Programs
HM Treasury	Her Majesty's Treasure
HR	Human Resources
IE	Impact Evaluation
INTOSAI	International Organization of Supreme Audit Institutions
LMs	Line Ministries
M&E	Monitoring and Evaluation
MSD	Ministry of Social Development
MoF	Ministry of Finance
MTEF	Medium Term Expenditure Framework
NGO	Non-Governmental Organisation
OECD	Organisation for Economic Co-operation and Development
OSGP	Office of the Secretary-General of the Presidency
SAI	Supreme Audit Institution
SECG	Sistema de Evaluación y Control de la Gestión

Executive Summary

The evaluation process developed by the Chilean Budget Directorate (DIPRES) represents an example of international good practice in the design and execution of *ex-ante* and *ex-post* evaluations. Chile has developed a comprehensive set of evaluation tools, ranging from *ex-ante* appraisal of new spending proposals to *ex-post* evaluation of individual programmes and institutional evaluations covering multiple programmes of a ministry.

Chile has placed a strong emphasis on *ex-ante* and *ex-post* evaluation as tools for performance management. This contrasts with other OECD countries that place more reliance on other tools, such as programme and performance budgeting and spending reviews to improve the quality of public spending.

Ex-ante analysis of new spending proposals, following a well-developed methodology and involving co-operation between DIPRES, the Ministry of Social Development (MSD) and Line Ministries (LMs), improves the quality of spending, through systematic analysis of programme design, including use of logical frameworks and indicators that also creates a strong basis for programme monitoring and evaluation. However, there is no clear connection between the *ex-ante* and the *ex-post* evaluation systems, especially in the case of social programmes, where the evaluation function is split between the MSD and DIPRES.

Chile has a well-developed *ex-post* evaluation system. However, after 20 years of implementation, there is widespread disappointment about the limited impact it has had on budget allocations, programme redesign, or closure of poorly-performing programmes. The overall programme of *ex-post* evaluations averages less than 5% of total government expenditures annually and most individual evaluations are also small in terms of expenditure coverage. Programmes in Chile are numerous but mostly small in scale compared to other OECD countries.

At the same time the range of issues covered by evaluation is broad, extending beyond financial performance to consider issues of policy and programme design, programme management and overall impact. This holistic approach could be considered a strength of the evaluation system, but it has also given rise to concerns about whether this scope goes beyond the legitimate interests of DIPRES.

At an annual cost of USD 1 million, the evaluation system is not expensive and yields benefits. However, it is possible that the same money, spread across a broader range of tools, could have a bigger impact on the quality of public spending. For example, spending reviews have been extensively used in OECD countries with positive results both in terms of savings and reallocation of resources to match spending priorities.

The main responsibility for the evaluation function rests with the budget department of the Ministry of Finance (DIPRES). The MSD also plays a significant role, in respect of *ex-ante* evaluation (appraisal) of social programmes. Congress is involved in deciding with DIPRES, which areas of expenditure should be subject to evaluation each year. Congress also reviews all evaluation reports although albeit with limited analytical capacity and with limited, detailed follow up of evaluation findings.

Concern about DIPRES's dominance of the evaluation function, together with issues of quality and the apparent lack of impact, has given rise to a long running debate about the governance of the evaluation function. Successive governments have made commitments

to set up an independent evaluation agency but no action has been taken. The experience of other OECD countries provides no best practice model for governance of the evaluation function. However, it does suggest potential areas for improvement, including greater participation of LMs and changes in the process for scrutiny of evaluations.

A fundamental obstacle to evaluations having budgetary impact is the current structure of the budget. The budget does not have a programmatic structure and programmes appear either as a single line item, or are hidden within a line item, (for example, transfers). This makes it hard to track spending or evaluate programme efficiency and effectiveness. While efforts have been made to prepare spending reports on programme execution, these have to be prepared manually, and represent a best estimate of expenditures, rather than audited expenditures. This inevitably reduces the budgetary relevance of evaluation findings. This contrasts with other OECD countries where budgets are organised and appropriated on a programmatic basis, enabling governments to better track programme expenditure and relate this to performance.

The quality of *ex-post* evaluations is not consistent. Evaluations are carried out by independent panels of experts, following clear terms of reference. However, the process of panel selection means that panel composition is unpredictable and the quality of work is sometimes affected by ineffective team leadership and co-ordination. Standardised budgets and timeframes for evaluations also mean that more complex programmes cannot be examined in the same depth as simpler programmes, sometimes resulting in loss of credibility and impact.

The feedback loop, whereby evaluation findings influence subsequent budget decisions is weak. In part this is due to inherent problems in linking budget allocations to outputs and outcomes, caused by the complexity of the problems that governments deal with and the often unclear relationships between financial inputs, outputs and outcomes. In part however it can also be attributed to poor internal co-ordination within DIPRES, between the two departments that deal with evaluation and the budget formulation, and limited follow up with LMs.

Evaluation reports are published and readily accessible on the website of DIPRES thus ensuring a basic level of transparency. However, the level of interest and engagement of politicians and representatives of civil society is low. In this respect, Chile is no different from most other OECD countries. The press shows some interest in evaluation findings, and regularly report on them but journalists are generally not sufficiently knowledgeable enough to provide good quality critical analysis.

Recommendations

Chile has invested significant resources in developing evaluation as a key tool for promoting the effectiveness of public expenditures. A high level recommendation is that Chile broadens the range of analytical tools that it uses to assess spending effectiveness, efficiency and policy alignment. New tools should include programme and performance budgeting, with an initial focus on adopting a programmatic budget classification; interim monitoring and review of programmes; and spending reviews of programmes, groups of programmes or sectors. Programme budgeting would facilitate expenditure tracking and analysis, and so provide a much stronger basis for budgetary analysis as part of the evaluation process. Interim performance monitoring and evaluation would support problem identification and course correction during implementation, thus reducing the chance of

poor outcomes. Spending reviews have proven effective at addressing issues of misalignment between existing spending allocations and government policy priorities.

There is a long standing, unresolved debate about the independence of the evaluation function and its positioning within Government. OECD recommends that leadership of the *ex-post* evaluation function remains within DIPRES, but that responsibility for evaluations is partly devolved to other LMs with the objective of improving quality and ownership and thereby impacting evaluations as well as overall coverage. In a revised set-up DIPRES would retain responsibility for evaluation methodology and oversight of evaluation efforts across government and would carry out evaluations of important programmes focusing on effectiveness, economy and efficiency. LMs would have a mandate to initiate their own programme evaluations, following central guidance, and focused on the effectiveness of policies and programme management. In this way the total coverage of could also potentially be increased from the current level of around 5% a year.

Congress's role in the selection of evaluation subjects and scrutiny of evaluation should be enhanced, both through training and capacity building of the Congressional secretariat. Congress's influence could be further strengthened if the CGR could begin carrying out performance audits, with their work programme responding to specific requests from Congress.

Ex-ante and *ex-post* evaluation should also be linked more closely through active management by DIPRES, establishing a common database and bringing together, at critical moments in the evaluation and budget decision making cycles, officials from the two departments in DIPRES responsible for evaluation and budget formulation, plus co-ordination mechanisms with MSD and other agencies.

Chile should consider more structured engagement with civil society organisations. A possible mechanism for this would be a Civil Society Council, whose role would be to independently review and comment on budget matters including the proposed programme of evaluations and findings. The establishment of such councils is already provided for under Chilean law, although none has yet been established for DIPRES.

Linked to the proposed revamp of the evaluation system, the centre of government should make explicit communication efforts designed to emphasise political support for programme evaluation, and its role in promoting effective and accountable management of public spending. This, combined with other recommendations, would help ensure that there is effective co-operation and follow-up by LMs.

Table of contents

Acronyms	2
Executive Summary	3
Recommendations.....	4
1. Context and institutional set-up.....	9
Legal framework.....	9
Institutional roles and responsibilities	9
An ongoing debate on the institutional position of the evaluation function	16
Assessment and future orientations for reform	20
2. Evaluation system methodology and coverage	22
Chile’s monitoring and evaluation system.....	22
A fragmented system for <i>ex-ante</i> evaluation	23
A well-established <i>ex-post</i> evaluation system	24
Potential use of interim evaluations	27
Selection of programmes to be evaluated	28
Key methodological challenges	29
Limited evaluation coverage.....	31
Assessment and future orientations for reform	34
3. Evaluation results and impact.....	36
A formal process to follow-up on evaluation results	36
Low perceived impact on both resource allocation and policy.....	36
Key challenges limiting the impact of evaluation results	40
Lack of specific tools to increase efficiency and improve budget allocation	43
4. Summary findings and recommendations.....	52
References.....	54

Figures

Figure 1. DIPRES organizational chart.....	10
Figure 2. DIPRES Public Management Control System.....	11
Figure 3. Governance of Evaluations.....	17
Figure 4. <i>Ex-post</i> evaluation process.....	25
Figure 5. Organisations responsible for deciding what will be evaluated.....	29
Figure 6. Programmes and institutions evaluated by evaluation streams.....	31
Figure 7. Coverage of Programmes and Institutions Evaluated in terms of budget.....	32
Figure 8. Coverage of Programmes Evaluated by Ministry	32
Figure 9. Characteristics of evaluation results.....	37
Figure 10. Utility level of the evaluation results	38
Figure 3. Evaluation results by performance category	39
Figure 12. Impact of evaluation processes on budget decision making in OECD countries.....	40

Figure 13. General trends towards increased use of all types of spending reviews 45
 Figure 6. Primarily responsible government actor for spending review procedures (2018) 47
 Figure 15. Application of spending reviews (2018)..... 48

Boxes

Box 1. UK - House of Commons Scrutiny Unit..... 13
 Box 2. Audit of non-financial performance information by the Auditor-General of South Africa..... 14
 Box 3. Canada M&E Institutional Arrangement 17
 Box 4. CONEVAL: Mexican Independent Government Agency 18
 Box 5. UK: What Works Network 19
 Box 6. Monitoring system in Chile 22
 Box 7. United States - Quarterly performance review 27
 Box 3. Canada: Led performance and impact assessment of conducting horizontal reviews 33
 Box 9. UK and Ireland. In-house evaluation capability 35
 Box 10. French programme budgeting system..... 42
 Box 11. Denmark: Special Studies review 46
 Box 12. Spending Review process and main steps 49

Introduction

As part of the discussions on the 2018 Budget Law, the Chilean Congress and the government agreed to review and assess the Program for the Evaluation of Programs and Institutions which is part of the Management, Evaluation and Control System of the Budget Directorate (DIPRES) of the Ministry of Finance (MoF). The OECD was invited to conduct this review, which was carried out between July 2018 and February 2019.

The purpose of this review is to analyse the current institutional arrangements, evaluation instruments and processes involved in the evaluation of spending programmes and institutions, and follow-up actions. In addition, the review examines the extent to which DIPRES evaluation evidence is used to inform budget decision making, as well as the wider relevance and use of evidence generated by evaluations. In particular, it looks at the extent to which evaluations provide evidence of whether programmes are meeting their objectives and whether evaluations include the right elements and evidence to justify the continuation, modification or closure of programmes.

The review also considers evidence of the extent to which the resources used for evaluation have contributed to better use in public resources, including efficiency gains in the programmes evaluated and/or changes in resource allocation (at government level and among programmes) and whether the DIPRES evaluation system represents an effective use of resources. It is important to note that the review does not provide a quantitative assessment of the impact of the evaluation system, but rather assesses and provides recommendations in respect of its qualitative impact on resource allocation and policymaking and the governance framework of the evaluation system.

The report is divided in three parts, the first chapter focuses on describing the institutional arrangement that supports and reinforces the Chilean evaluation system, namely the legal framework, the main actors, roles and responsibilities, and the debate on where the evaluation function should be located in terms of institutional design. The second chapter, analyses the evaluation system in detail, including policies and strategic approaches to *ex-ante* and *ex-post* evaluations, the procedure and criteria used to prioritise the programmes to be evaluated, some identified methodological challenges, and an assessment of the coverage of the evaluation system in terms of programmes and institutions evaluated. Finally, the third chapter assesses the impact of evaluation results in both resource allocation and policy making, and explores the main limitations and constraints on use of evaluation results to inform decision making.

1. Context and institutional set-up

After 20 years of implementation, the evaluation system is embedded in a robust institutional and legal framework. This comprises a set of formal rules and regulations, institutional roles and responsibilities and internal organisational and management arrangements that ensure the smooth functioning of the system. This chapter presents and discusses each of these three elements and offers a substantive assessment of its suitability and tailored recommendations.

The institutional set-up of the evaluation function has been the subject of a long running debate in Chile, driven by concerns that locating the function in DIPRES may reduce independence and objectivity, and by concerns about the limited impact of evaluation on policy making and budget allocations. This issue is also discussed in this section, together with possible alternative models from other OECD countries.

Legal framework

Chile has a well-developed legal framework that defines key roles and responsibilities as well as the principles and characteristics of the different evaluation streams. The basis for the *ex-post* evaluation system of public programmes is set in article 52 of the Organic Law of Financial Administration of the State. The main objective of the system is to evaluate the efficiency and effectiveness of public programmes and institutions, and provide information to support management and assessment for efficient decision making in the allocation of public resources.

A more detailed regulatory framework was provided in 2003 with the issuance of Decree No. 1177. This regulation includes an explanation of the different *ex-post* evaluation streams and their use, a set of guiding principles based on independence and transparency, and the definition of key roles and responsibilities. This gives the Budget Office (DIPRES) the leading role, guiding and co-ordinating the implementation of the different evaluation streams. It also prescribes that evaluation should be carried out by authorities and institutions with specific functions that must meet the principles of technical reliability, relevance and opportunity.

In 2011, the Law No. 20.530 created the Ministry of Social Development (MSD). This Law assigned responsibility for all policies, plans and programmes in the area of equity and/or social development to this ministry, including the function of evaluating all new or significant reformulated social programmes. It also provides key steps for the evaluation process and creates the Integrated Bank of Social Programs (BIPS).

Both the MSD and DIPRES have developed specific regulations and guidelines to define the process, criteria and objectives of each of the evaluation streams. In particular, in 2015 DIPRES issued conceptual and methodological guidelines for the *ex-post* evaluation system, updating and unifying previous manuals in a single document.

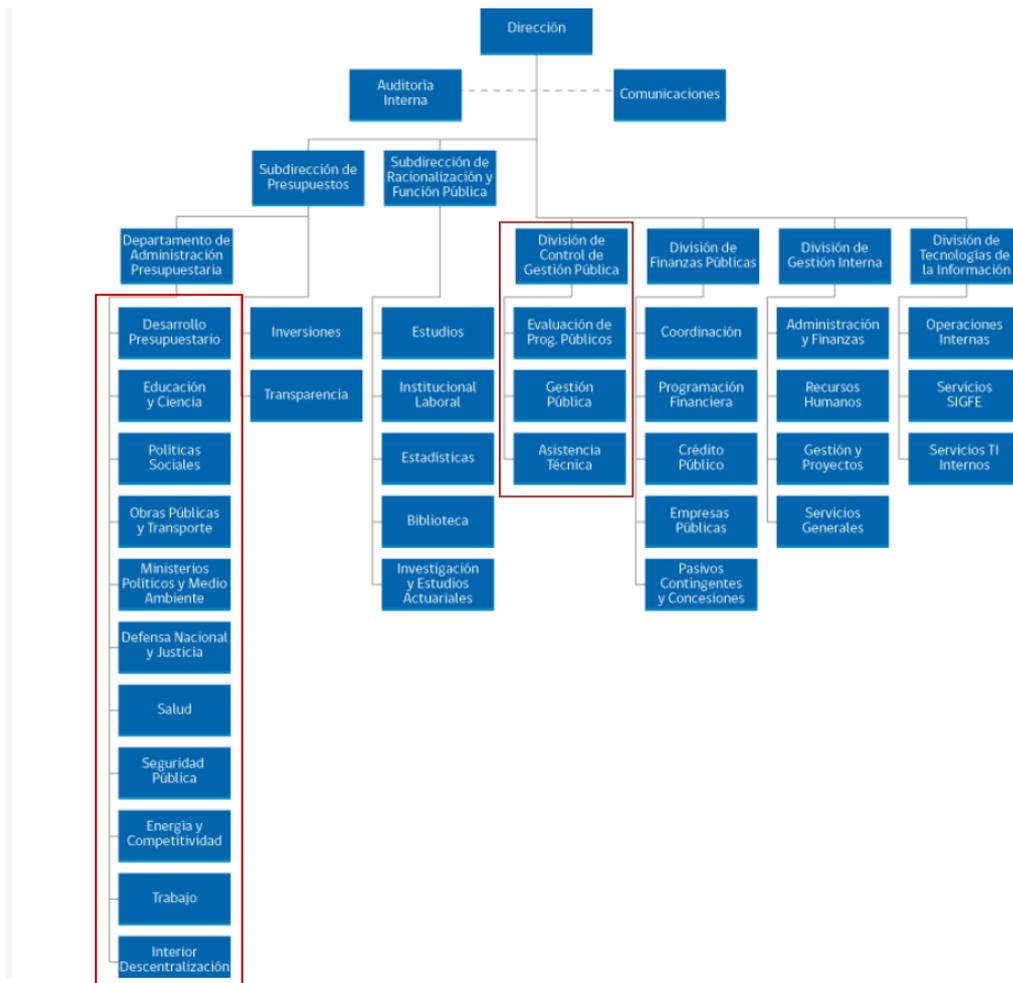
Institutional roles and responsibilities

Budget Office

DIPRES is the leading agency in the ongoing process of improving the quality of public spending in Chile. Its main objectives are to strengthen the budget as a mechanism to reach

the most efficient allocation of public funding in line with government priorities; foresee public revenues and expenditures in order to ensure fiscal sustainability in the medium and long term; and reinforce multi-level co-ordination as a way to streamline the budget bill. In order to do this, DIPRES has a staff of around 200 people. In addition to the classic structure of having sectoral divisions shadowing respective spending ministries, the Budget Office has special units dealing with performance and results information, long-term budget studies, and financial management issues (Figure 1).

Figure 1. DIPRES organizational chart



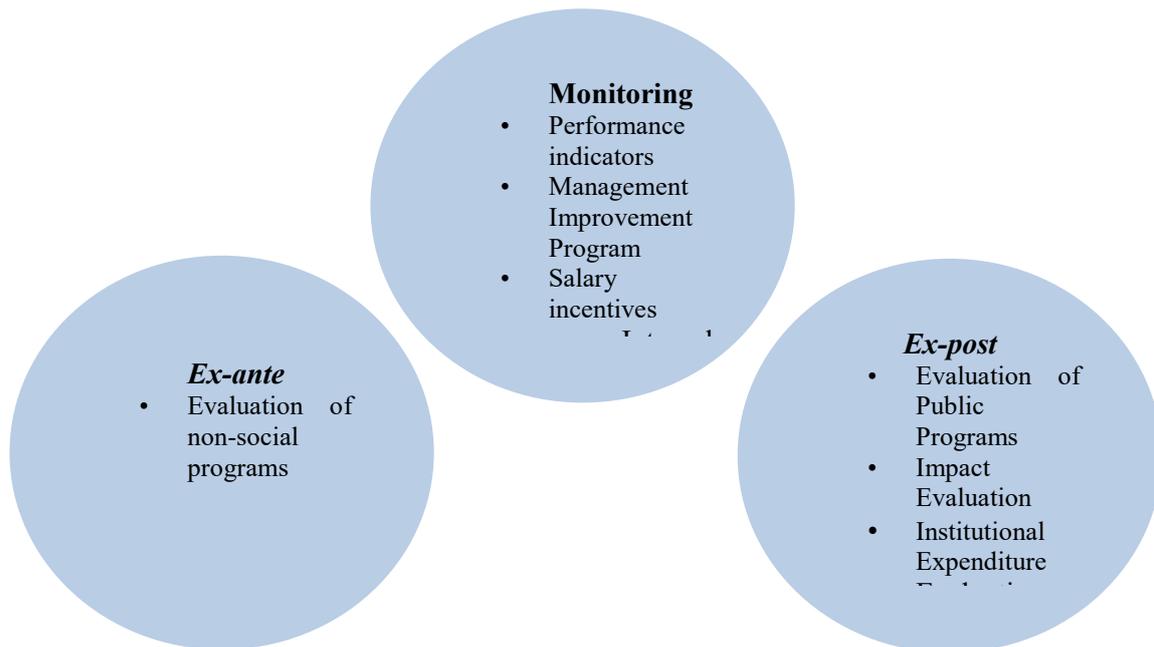
Source: DIPRES

The Budget Sub-Directorate deals with the budget preparation process and has two units: the Budget Management Department and the Investments and Transparency Department. The former is mainly focused on the preparation of the budget bill and there is a public servant in charge of budget formulation for each policy sector (policy sector experts). The latter, is in charge of the annual investment plan and the national transparency strategy.

The evaluation function is located in a separate Division of DIPRES, namely the Public Management Control Division. This Division works in consultation with the Inter-ministerial Committee and the National Congress, to define the evaluation pipeline (see

below). This division is also responsible for developing the evaluation methodology, defining the organisation and financing of the evaluation work programme, reviewing and approving the evaluation reports (quality control) and sending the evaluation reports to Congress and other public institutions. Besides the *ex-post* evaluation function, this division of DIPRES is responsible for the *ex-ante* evaluation of non-social programmes, the definition of performance indicators, and the management of the remuneration incentive programme (Management Improvement Programme), among other activities (Figure 2).

Figure 2. DIPRES Public Management Control System



Source: Authors

Inter-ministerial Committee

The evaluation process is guided by an Inter-ministerial committee, comprising of representatives of DIPRES, the Ministry of the General-Secretary of the Presidency, the Ministry of Social Development and the Ministry of Finance, in consultation with the Congress. This committee is responsible for guaranteeing that the evaluation programme is consistent with government policies and priorities; the conclusions of evaluations, alongside technical support, are available to stakeholders; and co-ordination is encouraged to ensure the quality of the work, including during the selection of consultants/advisors. Every year there are two key moments for the Inter-ministerial Committee. Firstly, before the beginning of the budget bill discussion (between August and September), the Committee is in charge of presenting the general results from the evaluations made that year to Congress. Secondly, between September and December, the Committee is responsible for preparing the proposal for forthcoming evaluations in order for it to be discussed and approved by the Congress.

Evaluators

The evaluations are carried out by evaluators from outside the public sector, typically technical experts from universities, consultancy firms or international organisations, with competence in the area of evaluation, and selected by public tendering. Evaluators (*Panel Evaluador*) are responsible for organising and analysing the information provided by related ministries and public agencies, and work in co-ordination with Line Ministries (LMs) and the DIPRES in order to have access to the information and research necessary to fulfil the evaluation. Finally, they should also be available to attend the aforementioned Congress meetings, if convened.

Ministry of Social Development

Since 2011, the Ministry of Social Development (MSD) has assumed responsibility for *ex-ante* evaluation of all social programmes and public investment initiatives. Locating this responsibility in the MSD is intended to improve the co-ordination and coherence of social investment across all government agencies. The co-ordination process is also supported by an Inter-ministerial Committee for Social Development. Within the MSD, the Sub-Directorate of Social Evaluation is in charge of the evaluation function. Evaluation covers both new social programmes and those that require significant reformulation.

Line ministries and agencies

In comparison with other OECD countries line ministries in Chile play a very limited role in the evaluation process, playing no role in the selection of the programmes to be evaluated having no responsibility for the conduct of evaluations beyond providing information to the evaluation team in the form of written documents, data and face to face interviews. In contrast in the UK evaluations are produced within departments by in-house analysts, either in a distinct evaluation unit, or as part of the financial management function. Some evaluation work is also tendered out to external researchers. The Treasury's (Ministry of Finance) is responsible for providing guidance on evaluation which is in the form of the "Magenta Book". The book is divided into two parts. The first part is designed for policy-makers and sets out the benefits of evaluation and steps that people can take to make sure their policy will produce good quality results for evaluation. The second part is more technical and goes into the details of planning and undertaking an evaluation and how to answer evaluation research questions using different approaches

Once the preliminary findings have been established, line ministries and agencies review and comment on the findings to be delivered to the Ministry of Finance and propose follow up actions, making institutional commitments to DIPRES. Following this, they are responsible for reporting to the Budget Office on the implementation of the institutional commitments every six months.

Congress

The National Congress of Chile has a formal role in determining which programmes should be evaluated, and in reviewing the findings of evaluations. During the discussion of the budget bill each year the Special Budget Commission decides which programmes should be evaluated based on a proposal from the Executive. This decision is embodied in a protocol of agreement between the Ministry of Finance and the Congress, or is established by an exempt resolution of the Ministry of Finance.

All final evaluation reports are made available to the Congress and the public, and their summaries are included in the budget information papers in the form of “executive minutes”. Congress also reviews all the results of evaluations, including indicators of performance and achievement of goals, as part of the annual budget package presented during the budget approval phase.

Despite this formal involvement in the process, in practice Congress’s influence is limited and members of Congress feel largely excluded from the budgetary processes. This is partly due to deep seated government concerns that ceding more authority to Congress could undermine Chile’s hard won achievements in fiscal stability. In exercising its oversight role, one of the main constraints members face is limited capacity to understand and analyse the budget. The budget document is long, detailed and technical and legislators rely on experts in a small Budget Advisory Unit. The small size of the Budget Advisory Unit allows only limited scope and depth of analysis, and to some extent, a dispersion of the topics analysed. Among other OECD countries, the UK offers an example of effective support to parliament in carrying out its scrutiny role (Box 1).

Box 1. UK - House of Commons Scrutiny Unit

The House of Commons’ Scrutiny Unit supports select committees in examining the expenditure and performance of government, and the relationships between spending and delivery of outcomes. It does this by promoting the value of linking examination of spending with examination of outcomes, by helping committees analyse spending patterns alongside performance and by pressing the government to improve the information available and promoting Parliament’s interests of holding the executive to account. For instance, the Scrutiny Unit has:

- produced a guide to committees on Better Financial Scrutiny which encourages examination of spending and outcomes throughout a programme’s lifespan, and sets out good practice;
- contributed financial and performance material to committee inquiries, including briefings, questions, reports, and analysis of impact assessments;
- analysed and briefed committees both on spending and trends in performance, using published indicators, when committees hold their hearings with Ministers on Government departments’ annual reports and accounts;
- followed up a committee recommendation for government departments to produce annual mid-year reports; and
- engaged with government in developing proposals to improve and simplify Government accounts for the benefit of Parliamentary users.

Congress has not demonstrated strong or sustained interest in the findings of *ex-post* evaluations. Due to the structure of the budget, Congress also faces difficulties linking programme evaluation findings to budget allocations due to the fact that budgetary programmes do not correspond to operational programmes (see Chapter 3). This reduces their capacity to influence budget decisions, deliberate them in a meaningful manner and hold the government accountable for performance.

Civil Society

Evaluations are publicly available on the DIPRES website. However, there is little engagement with civil society organisations in respect of evaluation findings. There is potential for greater involvement from a number of NGOs, both those that take an interest in budgetary issues, such as the *Observatorio Fiscal*, and those with a focus on specific policy areas subject to evaluation, such as environment, health, education, and social welfare. Civil society organisations' ability to engage in legislative debate on budgetary issues is constrained by a lack of relevant information related to individual programmes, given that the budget and accounts do not report allocations or expenditure on a programmatic basis.

Supreme Audit Institution

The Office of the Comptroller General of the Republic (*Controlaría General de la República*, CGR), which is the supreme audit institution of Chile, is not involved in reviewing programme performance. It is understood that a key reason for this is that the statute governing the work of the CGR excludes them from questioning government policy decisions. This has been interpreted as preventing their involvement in reviewing programme effectiveness and performance and their work is therefore focused on compliance audit. The absence of a performance audit function within the CGR makes the current evaluation function of DIPRES that much more important.

This contrasts with many OECD countries where SAI's examine issues of economy, efficiency and effectiveness in use of government funds, as per INTOSAI standards and also audit performance information reported in Government financial statements. For example, the Auditor-General of South Africa has audited non-financial performance information since 2005-2006 (Box 2).

Box 2. Audit of non-financial performance information by the Auditor-General of South Africa

Beginning in 2005-06, the Auditor-General of South Africa (AGSA) began to audit non-financial performance information. This stemmed from the requirement under the Public Finance Management Act, 1999 (Act 1/1999) and the Municipal Finance Management Act, 2003 (Act 56/2003) that accounting officers must report annually on the performance against predetermined objectives.

The AGSA audit approach focuses on:

- Understanding and testing the policies, procedures and controls related to the management of performance information.
- Understanding and testing systems and controls relevant to collecting, monitoring and reporting performance information
- Confirming the existence and consistency of performance information between the strategic/annual performance plan, quarterly reports and annual performance report and its presentation in annual reports
- Comparing reported performance information to relevant source documentation and auditing the validity, accuracy and completeness of reported information.

The AGSA audit criteria includes:

- 1) Compliance with planning and reporting requirements
 - Existence of defined objectives, indicators and targets, with information reported against those predetermined objectives, indicators and targets
 - Timeliness of the reporting of performance information within two months after the end of the year
 - Presentation of performance information using the National Treasury guidelines; the consistency of performance information in tables and the narrative of the annual report; with clear explanations of material differences between actual and planned performance
- 2) Usefulness of reported performance information
 - Measurability of objectives using indicators and targets, with well-defined indicators and SMART (specific, measurable, attainable, relevant and time bound) targets
 - Relevance with a clear and logical link between the objectives, outcomes, outputs, indicators and performance targets
 - Consistency between objectives, indicators and targets and between various planning and reporting documents
- 3) Reliability of reported performance information
 - Validity of actual reported performance
 - Accuracy in the amounts, numbers and other data relating to reported performance
 - Completeness of results and events in the annual performance report

Source: (OECD, 2014^[1])

Increasingly, national audit offices are using performance audit to provide insights into complex problems and risks, such as modernising outdated financial regulatory systems and protecting public safety. In the view of the US Government Accountability Office, for example, performance audit aims to provide objective analysis that management and those charged with governance and oversight can use to improve programme performance and operations, reduce costs, facilitate decision making and contribute to public accountability. SAIs often separate performance auditing from financial and compliance auditing, with personnel working on performance auditing having different backgrounds and skills from those selected for the other audit streams.

As highlighted in the previous section, Chile's SAI does not currently perform an evaluation function. In articulating its reasons for not conducting performance audit, the CGR cites the restriction on auditing the "merit" of political or administrative decisions, and the fact that the verification and assessment of compliance with the objectives and goals is considered as the mandate of the administration (OECD, 2014^[1]) As highlighted in the OECD Review of Chile's Supreme Audit Institution (2014) performance auditing does not need to question the merit of intentions and decisions; instead, it may focus on examining whether possible shortcomings in organisation, management and support

systems have affected the capacity to deliver the expected results. Likewise, any role the SAI has could complement and strengthen - rather than duplicate - the work of the central administration in programme evaluation.

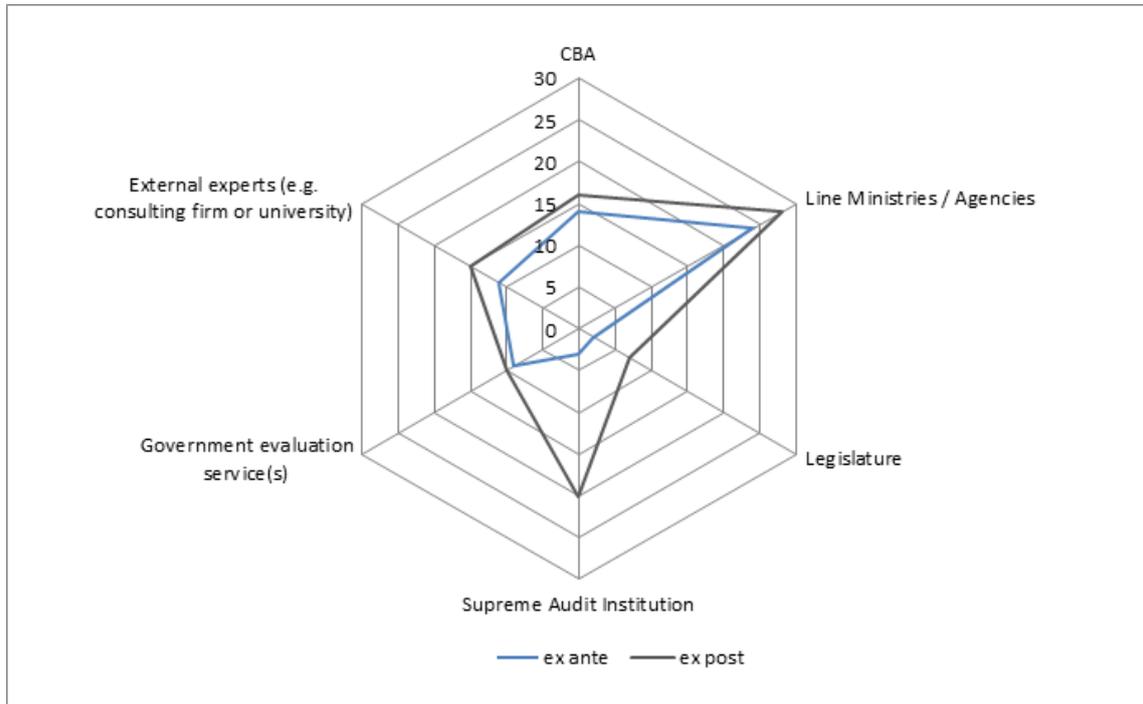
An ongoing debate on the institutional position of the evaluation function

The institutional positioning of the evaluation function has been a subject of debate over a long period of time. The debate has been prompted by concern about DIPRES's dominance of the evaluation function, together with issues of quality and the perceived low impact. The current set-up raises concerns in respect of objectivity and influence, given DIPRES's role in approving and controlling expenditure on the one hand and in evaluating programmes on the other hand.

Proposals to create a new institutional framework with an independent agency or directorate responsible for evaluation has been part of the agenda of the last two administrations and was a topic discussed during the last presidential elections. There have been different proposals from the executive and academia highlighting the roles and responsibilities that a potential new agency could undertake. It could, for example, assume responsibility for methodology and quality assurance of evaluations carried out by contractors. However, this model would be unusual; as most OECD countries have devolved the evaluation function to line ministries with methodological advice provided by the Ministry of Finance, Economy or the Centre of Government.

The experience of other OECD countries provides no clear best practice model for governance of the evaluation function. However, it is most common in OECD countries for the evaluation function to be devolved to line ministries, with the Ministry of Finance or another central agency providing standards and guidelines for the conduct of evaluations (Figure 3).

Figure 3. Governance of Evaluations



Source: OECD Performance Budgeting Survey 2018

This institutional set-up increases LMs’ buy-in and helps aligning the evaluation process with LMs’ needs and priorities. In Canada, for example, LMs are responsible for selecting, conducting and implementing the results of evaluations, under the guidance of the Treasury Board of Canada. Even though, under the new Policy on Results approach the Secretary of the Treasury Board of Canada has the ability to initiate evaluation exercises, the core evaluation function is devolved to LMs (Box 3).

Box 3. Canada M&E Institutional Arrangement

Canada has long used evaluation to support evidence-based policy development. For over 40 years, evaluation in various forms has been used to assess the performance of programmes operated by government departments. The 2006 Federal Accountability Act introduced a requirement that all grant and contribution programmes be evaluated every five years. The Policy on Results came into effect in 2016 and requires departments to have a five-year rolling departmental evaluation plan (Government of Canada, 2018).

Evaluation Units in each department undertake evaluation studies, but also have a broader role in relation to preparing evaluation planning reports and assessments, developing results-based management accountability frameworks and providing advice and training to programme managers on evaluation. Evaluations can also be conducted by contractors in whole, or in part, particularly when specific technical or subject matter expertise is required. Department managers use evaluation findings and recommendations to support policy and programme improvement, expenditure management, Cabinet decision making, and public reporting.

The Treasury Board of Canada Secretariat Results Division is the technical support unit for evaluations, setting standards, providing central leadership, guidance and support and using evaluation results, where appropriate, in decision making at the centre of government.

Source: (OECD, 2018^[2])

Few OECD Governments have a separate government evaluation service and where they exist they are attached to a larger ministry or agency, typically the Ministry of Finance the Ministry of Economy, the Ministry of Planning or the Centre of Government. Furthermore, all OECD countries with a separate government evaluation service effectively share responsibility for evaluations with line ministries, the SAI and/or the CBA. Independent government evaluation services can increase capacity and help ensure objectivity. However, there is a risk that the impact of evaluation results is weaker given their distance from the policymaking and resource allocation decision-making processes. The National Council for the Evaluation of Social Development Policy (CONEVAL) in Mexico illustrates some of these main advantages and limitations (Box 4).

Box 4. CONEVAL: Mexican Independent Government Agency

The National Council for the Evaluation of Social Development Policy (CONEVAL) is a decentralised agency of the Mexican Federal Government that co-ordinates the evaluation of the National Social Development Policy as well as other policies, programmes, or interventions related to social development. CONEVAL is the cornerstone of M&E system for the social sector in Mexico and is one of the leading agencies in the implementation of results-based management in the public sector more broadly. In particular it regulates, conducts, and co-ordinates the evaluation of social development policies at the national level; and develops guidelines and criteria for identifying and measuring poverty at the national, state, and municipal levels.

CONEVAL is a relatively independent government agency with a unique governance structure. Although it is located within the executive branch, and the federal government appoints its executive director, leadership of CONEVAL relies on an independent collegiate body made up of six academic councillors.

The expertise and specialisation of staff, and the participation of academics, has enabled CONEVAL to develop robust methods and to conduct and commission evaluations that are technically reliable and credible to both internal and external stakeholders. Furthermore, co-ordination and technical support has lowered the costs of evaluation practices in social agencies. However, due to its social sector focus CONEVAL has limited competences, increasing the technical capacity gap in evaluation between the social sector and other relevant areas and levels of the public administration.

CONEVAL's relations with ministries have also been difficult at times, particularly when evaluations point out implementation problems and programme shortfalls. Another important challenge is that being separate from the government's decision-making processes limits CONEVAL's capacity to enforce the use of evaluation information to directly influence policy and budget decisions (e.g., in budget and planning), even in the social sector.

Source: (Castro, Lopez-Acevedo and Busjeet, 2009^[3])

In Chile, there are some clear advantages from DIPRES's close involvement in evaluation, notably its financial expertise and authority. This has helped to ensure that LMs pay attention to evaluation findings. However, as the impact of evaluation findings on resource allocation decisions has proven to be weak, the level of attention paid to the evaluation by LMs has weakened over time.

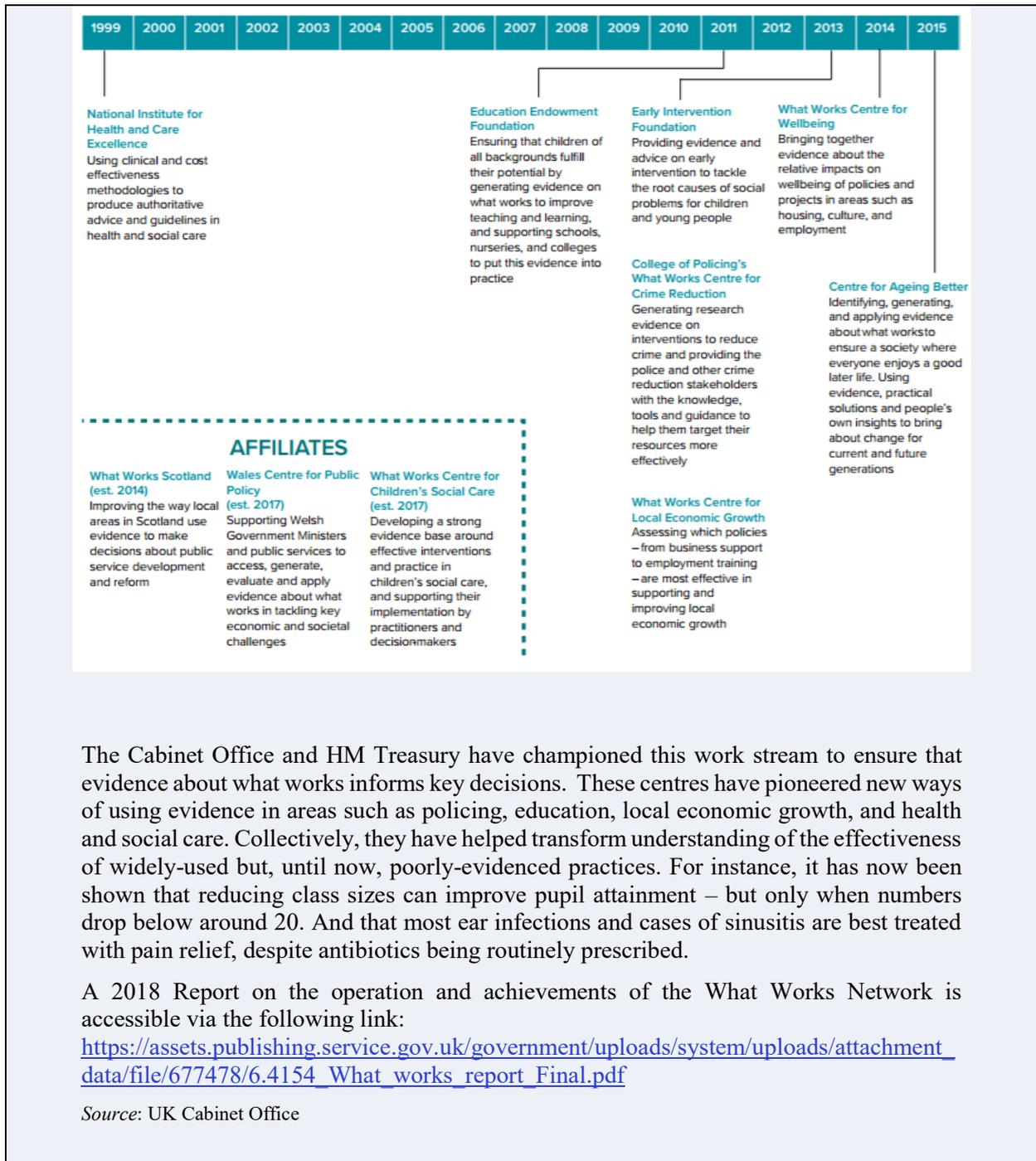
The current institutional arrangement, with the evaluation function located in DIPRES, helps to ensure the legitimacy of the evaluations results. Furthermore, the Inter-ministerial Committee that gathers the main stakeholders provides an institutional structure for developing the follow-up commitments and supporting changes. The Move being home of both the unit in charge of the public programmes evaluation and the Budget Management Department, allows an organisational environment in which evaluations results can feed into budget related decisions. However, these organisational arrangements do not necessarily facilitate an optimal communication flow between these two units and, in that sense, integration of evaluation results into budget decision-making.

The existing arrangement could be strengthened through consultation and participation by LMs, both in the selection of programmes and institutions to be evaluated and through involvement of technical experts from LMs as members of the evaluation teams. This could improve both the relevance, quality, and impact of evaluations.

Another option would be to develop a more informal network approach as has been developed by the UK. The What Works Network (see Box 5), supported by the centre of government (Cabinet Office) and the finance ministry (HM Treasury), encourages independent evidence-based analysis of policy and service delivery, making use of a network of independent institutions that specialise in analysis of public policy.

Box 5. UK: What Works Network

The What Works Network was launched in 2013 with one simple aim: to ensure that spending and delivery of public services is informed by the best available evidence. As of 2018, the Network now consists of 10 independent What Works Centres. Collectively, the centres have produced more than 280 evidence reviews in the last five years and commissioned or supported over 160 trials.



The Cabinet Office and HM Treasury have championed this work stream to ensure that evidence about what works informs key decisions. These centres have pioneered new ways of using evidence in areas such as policing, education, local economic growth, and health and social care. Collectively, they have helped transform understanding of the effectiveness of widely-used but, until now, poorly-evidenced practices. For instance, it has now been shown that reducing class sizes can improve pupil attainment – but only when numbers drop below around 20. And that most ear infections and cases of sinusitis are best treated with pain relief, despite antibiotics being routinely prescribed.

A 2018 Report on the operation and achievements of the What Works Network is accessible via the following link:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/677478/6.4154_What_works_report_Final.pdf

Source: UK Cabinet Office

Assessment and future orientations for reform

DIPRES plays a dominant role in the Chilean evaluation system. This reflects the centralised nature of the government in Chile and the authority of DIPRES, built on a track record of sound fiscal management and accumulated experience and expertise. This institutional design could bring some benefits; the process is taken seriously by LMs and evaluation findings can feed into resource allocation decisions. At the same time this model

presents a number of issues. First, Congress feels marginalised and questions the effectiveness of the current system. Even though it has a formal role in selecting the evaluation agenda and reviewing the results, in practice many congress members are not aware of the evaluation system and do not have the technical capacity to get involved in the process. Second, LMs have little influence over what is evaluated and are not directly involved in the evaluation process, which leads them to question both the relevance and the quality of evaluations. Third, political leaders have expressed concerns about the dual role of DIPRES, both as the evaluator and the final decision maker with respect to the budget, and suggested that the evaluation function should be transferred to an independent agency.

To address these concerns the following changes could be considered:

- Provide more support to Congress in the form of a strengthened secretariat with increased analytical capability, to **help members to engage more actively and critically in the evaluation process**. This would give greater effect to the provisions of the budget law.
- **Strengthen the role of line ministries in the evaluation process**. This would include involving LMs in the choice of programmes to be evaluated and including technical experts from the line ministries as members of evaluation teams. This should increase the relevance of subjects chosen for evaluation, improve the quality of evaluations and increase the likelihood that evaluation findings will be accepted and acted on by LMs.
- **A more fundamental reform would involve distributing responsibility for evaluation between DIPRES, LMs and the CGR, similar to the situation in other OECD countries**. In such an alternative model DIPRES-led evaluations could be more tightly focused on finance issues, namely efficiency, economy and effectiveness. LMs would become directly responsible for commissioning broader evaluations of their own programmes, with evaluation focused on policy design, impact and programme management. The CGR would establish a performance auditing function, carrying out audits on behalf of the Congress, within the limits of its governing statute. At a minimum, the scope of audit should cover the reliability of non-financial performance information that underpins programme management and accountability, but could extend well beyond this.
- On institutional positioning of the evaluation function, several options could be considered. DIPRES could continue to co-ordinate the evaluation process across government, including the development of evaluation methodology. This is similar to the situation in many OECD countries.
- **Strengthen the role of civil society in the evaluation process**. Civil society has had little involvement in the evaluation process, despite the reports being publicly available. In this respect Chile is no different to other OECD countries. Chile could consider more structured engagement with civil society organisations. A possible mechanism for this would be a Civil Society Council, whose role would be to independently review and comment on budget matters including the proposed programme of evaluations and findings. The establishment of such councils is already provided for under Chilean law, although none has yet been established for DIPRES.

2. Evaluation system methodology and coverage

After two decades of implementation Chile has a well-defined *ex-ante* and *ex-post* evaluation system that provides information on programme performance. In the process it has created an evaluation culture in the public sector and formed a critical mass of experienced evaluators. However, the system faces important challenges such as fragmentation across the several government units involved in process, opaque criteria for prioritising the evaluation agenda, and barriers to increasing the coverage of programme evaluations. This chapter examines the Chilean evaluation system methodology, the strategy behind it and the links that hold together various elements of the system. The chapter identifies key methodological challenges and ends with a summary assessment and recommendations.

Chile's monitoring and evaluation system

Chile has an advanced, robust and well-established system for obtaining information on agency and programme performance known as the Evaluation and Management Control System (*Sistema de Evaluación y Control de la Gestión*, SECG). This system aims to improve the effectiveness of policymaking and management throughout the central government, to create performance incentives for civil servants, and to make the budget results-oriented.

This report focuses mainly on DIPRES *ex-ante* and *ex-post* evaluation function. However, the SECG has other tools to monitor institutional performance, strategic priorities and specific programmes and projects (Box 6). Despite being implemented within the competences of the same Division, there are no explicit links between the different monitoring tools used by DIPRES and the *ex-ante* and *ex-post* evaluation system.

Box 6. Monitoring system in Chile

The Chilean monitoring system has three main actors: the MoF, the MSD, and the Office of the Secretary-General of the Presidency (OSGP). The following are the key monitoring tools used by these actors:

1. The H form (*Formulario H*): a document that accompanies the Budget Bill and comprises performance indicators that include qualitative information on public goods and services (MoF-DIPRES).
2. The Management Improvement Program (*Programa de Mejoramiento de la Gestión*) that grants allowance to public servants that reach specific targets (MoF-DIPRES).
3. Internal management indicators (*Indicadores de Gestión Interna*) that focus on internal processes and procedures (MoF-Plan ChileGestiona).
4. Collective performance agreements (*Convenios de desempeño colectivo*) that encourage team work within work units towards annual institutional targets (MoF-Others).
5. Individual performance agreements (*Convenios de desempeño individual*) that set strategic management targets for every civil servant (MoF-Civil Service).

6. Presidential priorities follow-up (OSGP-Presidencial Unit of Compliance Management - *Unidad Presidencial de Gestión del Cumplimiento*).
7. Social programmes follow-up: performance indicators of social programmes (MSD-Social Policy Division).

Source: (Irrarrázaval and De Los Ríos, 2014^[4])

A fragmented system for *ex-ante* evaluation

In contrast with the *ex-post* evaluation system that is fully centralised in DIPRES, the *ex-ante* evaluation function is divided between the Ministry of Social Development (MSD) and DIPRES.

Since its creation in 2011, the Sub-Directorate of Social Evaluation of the MSD has been responsible for evaluating and monitoring all new social programmes, including social evaluation of long-term capital investment projects. The *ex-ante* evaluation process uses a logical -framework approach, cost/benefit analysis and identification of performance indicators and targets. Other features of the process include checks on coherence with other programmes, identification of target beneficiary groups, and use of baseline surveys and random control trials to assess programme impact.

A similar *ex-ante* evaluation stream was created in DIPRES in 2009. This line assesses the design of each new non-social programme, establishing control groups for the evaluation, based on random trials, whenever possible. The evaluation process lasts 24 months. During this time DIPRES submit a preliminary report and a final evaluation report containing the results of the evaluation and recommendations.

Compared to other OECD countries, Chile's *ex-ante* evaluation process is comprehensive and well-regulated. What is involved in *ex-ante* assessment varies substantially in OECD countries. In many countries *ex-ante* assessment is understood to mean regulatory impact assessment, which may or may not examine the long-term budgetary impact of regulatory initiatives. In others, there are separate procedures for appraising long-term capital investment projects. In a few countries all of these are processes of *ex-ante* assessment and appraisal are covered by one set of guidelines (the "Green Book" in the case of the UK).

Despite, having robust methodologies and coherent regulation, the information system that holds the social programme evaluations is still basic, with a rigid structure that does not support the comparison of different programmes or measurement of progress in a timely manner. Furthermore, the MSD and DIPRES have different systems in place to report and monitor *ex-ante* evaluations. This arrangement is not common in OECD countries. Some countries like Austria have recognised the importance of having a unified portal to include all *ex ante* impact assessments. The Austrian Federal Performance Management Office and the Ministry of Finance have started a project to develop a web-based tool that would collect all *ex ante* impact assessments in a database and also digitise the relevant procedures.

The existence of a dual system has created some tension between MSD and DIPRES. The definition of a social programme is not always clear and the linkages between social and non-social programmes could be lost under the current system. Furthermore, there are some frictions between the MSD and DIPRES, regarding the *ex-post* evaluation function of social programmes. The MSD has expressed strong interest to build upon the experience gained

from performing the *ex-ante* evaluation and monitoring process, and would like to use this to undertake *ex-post* evaluation of social programmes. The function split between these two institutions imposes co-ordination challenges in the evaluation system that must be addressed to ensure there are links between the *ex-ante* and *ex-post* evaluation of social programmes, and that the methodologies, tools and criteria used for *ex-ante* evaluation are shared across the administration.

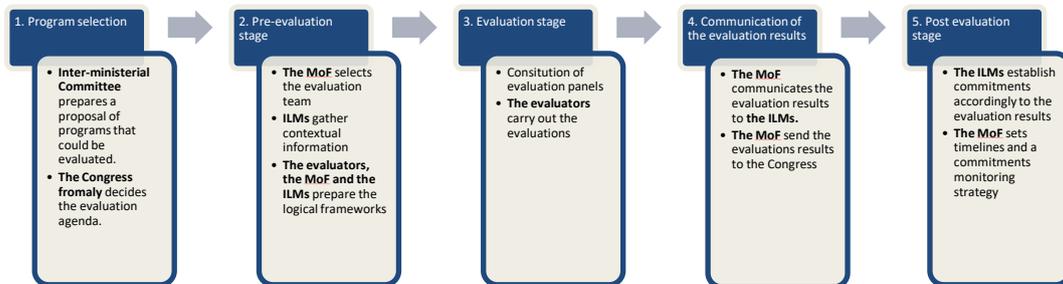
In contrast with the *ex-post* evaluation system, there is a direct link between the *ex-ante* evaluation process and the budget preparation process. Only programmes that have a positive *ex-ante* evaluation can be financed with the public budget. This system has helped improve the quality of programme design, ensuring minimum standards for both social and non-social programmes before resources are allocated. However, as it will be further explained below, this link is somehow lost after the programme is approved and starts to receive funding.

A well-established *ex-post* evaluation system

Chile has a well-defined *ex-post* evaluation function that is centrally regulated and systematically applied. Its main objective is to improve resource allocation and increase the quality of public spending. It started operating in 1997 with a pilot to evaluate the design and management of ongoing public programmes. In its experimental phase, the Evaluation of Public Programmes stream (EPG) was launched including the assessment of 20 public programmes. They responded to a number of commitments in the Budget Law concerning evaluation and transparency of public management. Substantial improvements have been carried out during these two decades, developing a robust framework, creating an evaluation culture in the public sector, and forming a critical mass of evaluators in the country.

Currently, the evaluation process is carried out in five main stages (Figure 4):

1. Selection of programs: in which the Inter-ministerial Committee prepares a proposal of the programmes that could be potentially evaluated. The proposal is submitted to Congress and is amended during the budget bill discussion.
2. Pre-evaluation stage: in which the evaluators are chosen, the contextual information of the programme is gathered, and the logical framework is prepared.
3. Evaluation stage: in which the evaluation teams (panels) are established and the evaluation is performed.
4. Communication of the evaluation results: in which the evaluation results are sent to the Congress and to the LMs.
5. Post-evaluation stage: in which the LMs use the evaluation results to define commitments, and the MoF sets timelines and a monitoring strategy.

Figure 4. *Ex-post* evaluation process

Source: Authors, based on DIPRES information

During the last 20 years, the *ex-post* evaluation system has gradually developed new streams of work with different scopes and objectives, but has always maintained the central focus of delivering performance information to strengthen the decision-making process and improve the quality of public spending. There are currently four main evaluation streams under DIPRES.

1. Evaluation of Public Programs (EPG)

In 1997, the Government of Chile started implementing evaluations of ongoing public programmes, aiming to improve public management and better inform the resource allocation process. EPGs uses a logical framework methodology to assess programme design, execution and reporting. The evaluation takes into account the following elements: 1) justification or main issue that the programme is aiming to address; 2) overall design; 3) organisation and management; and 4) results, including an efficiency, efficacy and economic assessment of the programme.

EPG are performed by panels of three independent experts, selected in a public tendering process. The evaluation process lasts for six months and has an average cost of USD 35 000. During this period, evaluators prepare preliminary and final reports providing the results of the evaluation, conclusions and recommendations. The timeline of the evaluation process is designed to ensure alignment with the budget formulation cycle.

2. Impact Evaluation (EI)

In 2001, the government created a new stream of evaluation that is designed to assess programme effectiveness based on impact measures. It looks at the causal relationship between programmes and results, using quasi-experimental methodologies. The analyses are based on extensive data collection, evaluations techniques such as cost-benefit and cost-effectiveness analysis, and control groups that allow the isolation of the effect of external factors.

Historically, IEs were performed by a consulting firm or a university selected by public tendering. However, last year DIPRES decided to build up its internal capacity and start implementing these evaluations in-house.

Impact assessments typically take 12 to 18 months and have an average cost of USD 200 000. During this period, preliminary reports are submitted, as well as a final report containing the results of the assessment and recommendations.

3. Institutional Expenditure Evaluation (EGI)

This evaluation stream was created in 2002 aiming to evaluate institutional performance. In contrast to other evaluation streams, the EGI does not evaluate single programmes. It focuses on the overall design, management, and results achieved by a particular public institution (i.e. a ministry or agency).

This evaluation stream uses a similar methodology to the one applied in the EPG stream, but it adapts the assessment requirements with a broader view, focusing on the strategic objectives, organisational structure and budget of the institution.

The evaluation process lasts approximately 8 months. During this time preliminary reports and a final evaluation report containing the evaluation results and recommendations are submitted.

4. Focused Scope Evaluation (EFA)

The latest evaluation stream was created in 2015, aiming to provide DIPRES with a less complex and faster evaluation process to assess relevant aspects of a particular programme. There are three possible focus areas:

1. Cost review: this looks at the efficiency with which a certain good or service is delivered. It aims to provide a unitary value that can be compared with a benchmark (market value), and explain the magnitude and causes of the difference between these two values.
2. Programme implementation: this looks at programme targeting, and the potential implementation gaps that exist between the initial programme design and actual execution.
3. Strategic design: this revises the foundations and information on which a programme is based in order to justify its existence, and the overall programme strategy.

EFA reports must be submitted within 4 months. Despite the short evaluation period, the stream lined process is not as fast as intended due to the lengthy procedure of selecting the independent evaluators. All four types of evaluation streams require selection of independent evaluators through a public tender process. DIPRES devotes a large amount of time and resources preparing and selecting external evaluators and consequently has limited capacity to undertake evaluations in-house.

Chile has a robust *ex-post* evaluation framework, with clear and defined product lines and a strong focus on issues of policy, programme design and programme management. The range of issues covered in *ex-post* evaluation varies widely across OECD countries. Certain countries including Canada, Japan and Spain mandate coverage of a comprehensive set of issues in their evaluations (OECD, Forthcoming^[5]). Despite the long-standing tradition and well developed framework, Chile could further improve the way the evaluation pipeline is

defined, as well as addressing other methodological challenges, as discussed in the following sections.

Potential use of interim evaluations

Interim evaluations could be a useful addition to the evaluation toolbox, enabling government to take corrective action in case of design or implementation weaknesses in a programme. This is relevant for all programmes and is especially important in the case of large capital investment programmes.

For example, in the United States Agency GPRAMA requires leaders and managers to use regular meetings, at least quarterly, to review data and drive progress toward key performance goals and other management-improvement priorities. For each APG, agency leaders conduct reviews at least quarterly to assess progress toward the goal, determine the risk of the goal not being met, and develop strategies to improve performance. There is research evidence that data-driven reviews, in which performance data are routinely discussed for management purposes are effective (Box 7).

Box 7. United States - Quarterly performance review

Another significant aspect of the US Government Performance and Results Modernization Act (2010) is that it requires formal routines for agency staff to discuss data. Agencies must hold quarterly reviews (sometimes called, data-driven reviews) of progress on agency priorities and other significant goals. The Chief Operating Officer is required to lead these reviews, and there is detailed discussion of progress on each goal by senior managers and the designated goal leader. The goal leader must track performance outcomes, understand why they rise and fall, and organise efforts for improvement.

- Review with the appropriate goal leader the progress achieved during the most recent quarter, overall trend data, and the likelihood of meeting the planned level of performance.
- Hold goal leaders accountable for knowing whether or not their performance indicators are trending in the right direction at a reasonable speed, and if they are not, for understanding why they are not and for having a plan to accelerate progress to the goal.
- Hold goal leaders accountable for knowing the quality of their data, for having a plan to improve it if necessary, and for filling critical evidence or other information gaps.
- Hold goal leaders accountable for identifying effective practices by searching the literature, looking for benchmarks, and analysing disaggregated data to find positive outliers among performance units.
- Hold goal leaders accountable for validating promising practices with replication demonstrations or other evidence-based methods.
- Review variations in performance trends across the organisation and delivery partners, identify possible reasons for each variance, and understand whether the variance points to promising practices or problems needing more attention.
- Include evaluation staff to share and review performance information and evaluation findings; better understand performance issues that evaluation and

research studies can help to address; and refine performance measures and indicators.

- Include, as appropriate, relevant personnel within and outside the agency who contribute to the accomplishment of each Agency Priority Goal (or other priority).
- Support the goal leaders in assuring other organisations and programmes are contributing as expected to Agency Priority Goals (or other priorities).
- Identify Agency Priority Goals (or other priorities) at risk of not achieving the planned level of performance and work with goal leaders to identify strategies that support performance improvement.
- Encourage a meaningful dialogue around what works, what does not, and the best way to move forward on the organisation's top priorities, using a variety of appropriate analytical and evaluation methods.
- Establish an environment that promotes learning and sharing openly about successes and challenges.
- Agree on follow-up actions at each meeting and track timely follow-through.

Source: The White House, Office of Management and Budget (2018).

In Canada, accounting officers have a personal legal obligation to appear before parliamentary committees and answer questions on the management responsibilities of their department. A similar arrangement for accounting officers exists in Ireland. Also in Ireland, an annual agreement is signed between a department head and their relevant Minister outlining performance objectives.

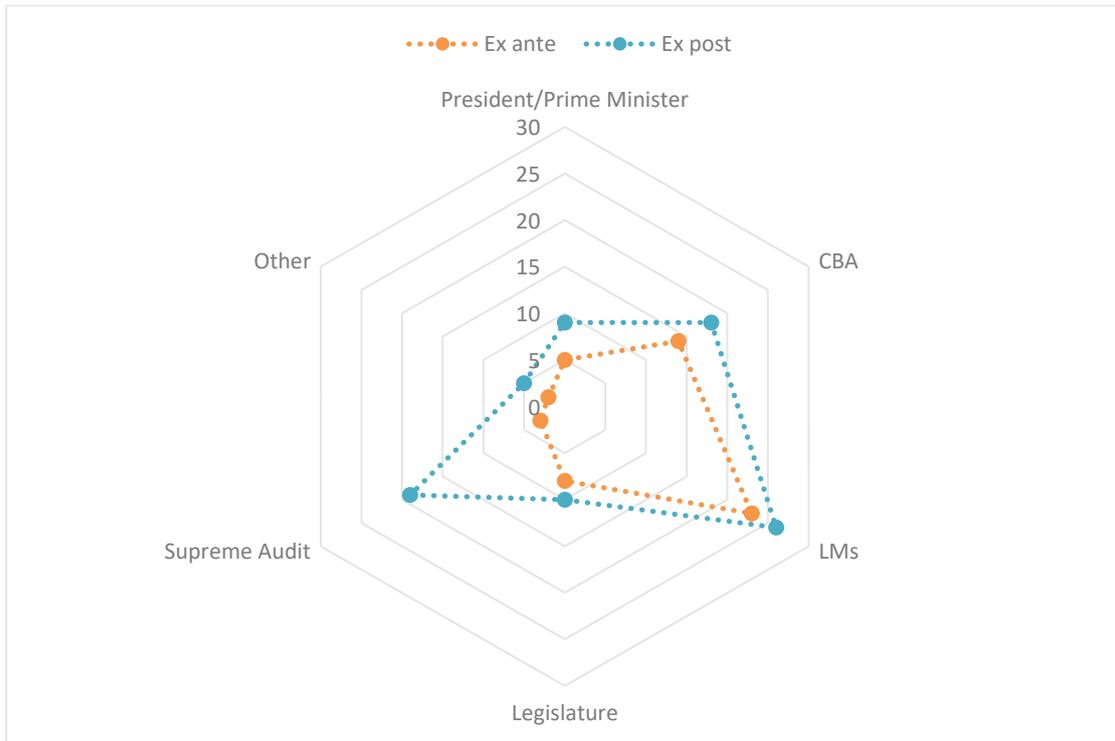
Selection of programmes to be evaluated

The current legal and regulatory framework does not stipulate any specific criteria for the selection of programmes to be evaluated during the calendar year. The selection process starts in September, when the annual evaluation results are presented to the Congress as a preamble of the discussion of the next budget bill. Between September and December, the Inter-ministerial Committee (IMC) prepares the proposal for next year's evaluations in order for it to be discussed and approved by Congress.

The list of programmes and institutions to be evaluated is set out in a Memorandum of Understanding (*Protocolo*) signed every year between the government and Congress during the budget approval process. This document is a formalised political agreement negotiated in parallel to the budget process. As such, the list is not defined based on predetermined criteria and the projects are not chosen based on budgetary or sectoral needs. As highlighted by some of the stakeholders interviewed, the selection process is often driven by the context, specific requirements from congressmen, the willingness of a ministry to co-operate in the evaluation, and other political dynamics.

In this sense, it is important to note that neither the LMs nor the policy experts from the Budget Management Department are officially involved in any of the stages of the selection process. This differs from the process in other OECD countries, where both SAI and LMs often have a prominent role in defining the evaluation pipeline (Figure 5).

Figure 5. Organisations responsible for deciding what will be evaluated



Source: 2018 OECD Performance Budgeting Survey

In practice, the IMC select programmes that meet certain basic criteria, such as those that have an M&E system in place, that have resources allocated, and have not had a recent evaluation. There are also other “informal” criteria that vary from year to year. In 2018, informal criteria included; linkage with presidential commitments, budgetary relevance, programmes with duplicity and/or complementarity issues, and programmes with performance challenges.

Key methodological challenges

As highlighted in the previous section, in Chile, the practice of evaluations is supported by clear methodological guidance and a well-developed process for the selection of expert panels and conduct of evaluations that has been developed and refined over a 20-year period. Despite this, the effectiveness of the evaluation process is adversely affected by a variety of issues as follows.

The unit of analysis is narrow

Except for Institutional Expenditure Evaluations, the Chilean system is mainly focused on the evaluation of specific individual programmes rather than on policy areas or government strategic goals. This has helped to develop very detailed and robust tools to look at one specific program, but it lacks a more comprehensive analysis that gives account of the way different common programmes are working towards achieving a common result. In character many programmes are more akin to small-scale policy initiatives. For example, three separate programmes were evaluated all aimed at keeping kids in school.

In the last couple of years DIPRES has tried to take some first steps towards more comprehensive analysis, commissioning the evaluation of two or more related programmes at the same time. However, without the proper guidelines and tools in place, the evaluation has been carried out as a group of evaluations rather than a comprehensive analysis on the way the different programmes interact. Evaluators have reported difficulties in carrying out such compound evaluations, since in practice, they need to evaluate each programme separately. Experienced evaluators reported that they are not presenting proposals for these compound evaluations because of the amount of work involved.

Evaluators also point to a problem of misalignment between what DIPRES understands to be a spending programme and line ministries' own understanding. As a result ministries and agencies often do not see the evaluation's findings as relevant. For example, money allocated in the budget to a "School achievement bonus programme" is understood by DIPRES as a programme, whereas for the line ministry this budget item is one component within a larger income support programme for families. As a result, line ministries do not always find the findings of evaluations relevant or useful. This problem is linked to the lack of programmatic structure in the budget that, if introduced, could ensure a common understanding of programmes (see Chapter 3).

Weaknesses in programme design

In a number of important sectors, such as health, new programmes are often developed without the benefit of logical frameworks, defined performance indicators or baseline numbers against which to conduct the evaluation. This undermines the basis of good evaluation and, as a result log frames and indicators are often imputed *ex-post*, as part of the evaluation process.

Co-ordination, teamwork and quality issues

Panels are separately constituted for each evaluation, with DIPRES deciding on the constitution of each panel, based on a public selection process. However, the effectiveness of leadership and teamwork appears to be variable. A panel co-ordinator is appointed for each evaluation but depending on the quality of the team as a whole, a disproportionate share of responsibility may fall on the co-ordinators. These factors lead to variability in the quality of evaluations.

Quality assurance

Evaluators interviewed by OECD commented that the quality assurance mechanism for evaluations was not clear. In practice this role lies with DIPRES. Evaluators highlighted DIPRES's commitment to ensure evaluation robustness and high quality. However, they also commented that comments on evaluation findings were sometimes excessive and too detailed, while at the same time DIPRES did not have sufficient understanding of the broader non-financial policy and implementation issues covered by the evaluations.

Collaboration with line ministries is becoming more difficult and access to information can be a challenge

Collaboration with LMs seems to become less effective over time. Evaluators reported that, compared to when the evaluation process was first introduced, ministries are now less inclined to make senior staff available to meet with the evaluators, depriving them of a strategic management perspective. Access to information has also been an issue with some LMs citing data privacy issues, for example in social programmes, as a reason not to share

important data, thereby impairing the quality of the evaluation. Lack of data and analysis can also lead to ambiguous recommendations.

Limited evaluation coverage

Despite concrete efforts to sustain evaluation capacities and the creation of new streams of work, evaluation coverage remains limited and varied across government. In the first 20 years of implementation (1997 to 2017), DIPRES carried out 403 evaluations, involving 560 programmes and institutions. EPG has been the evaluations stream most widely used, constituting almost 70% of the total number of evaluations carried out to date, followed by EI, which represents over 20% of total evaluations. There are fewer examples of the recently created EFA, representing just 2% of the total number of programmes and institutions evaluated (Figure 6).

**Figure 6. Programmes and institutions evaluated by evaluation streams
Absolute value (1997-2017)**



Note: Covers the evaluations completed, ongoing and those to be tendered.
Source: (Darville, Paula; Díaz, Rodrigo; Leiva, 2017^[6])

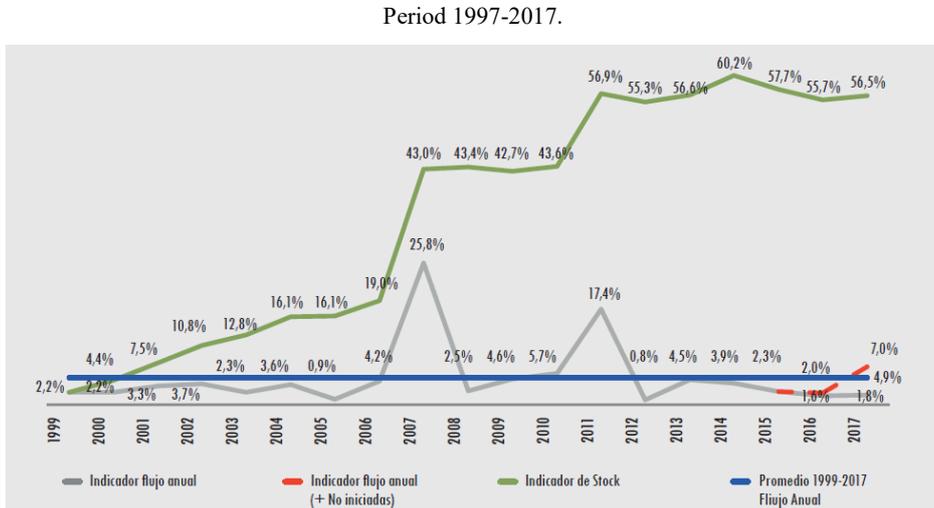
Despite the introduction of new evaluation streams, the number of EPG evaluations - the core of the evaluation system in Chile - has remained broadly stable, with an average of 18 evaluations per year between 1997 and 2017 and with increased use in recent years. Use of the EI stream has been more variable since its creation in 2001, with an average of 7 evaluations per year since 2001, but with declining use in recent years.

There are two main indicators of the low perceived coverage of the evaluation system: 1) the number of institution with programmes that have been subject to evaluation, and 2) the share of the national budget that has been evaluated. Since its implementation, the EGI stream has covered 23% of the total institutions of the Chilean government. Although cumulatively (stock indicator) 56.5% of the national budget has been covered since the evaluation system started (through the evaluation of programmes and institutions together), on average the annual coverage represents less than 5% of total spending. This figure is influenced by two important institutional evaluations¹

¹ Instituto de Normalización Previsional (INP) and Fondo Nacional de Salud (FONASA).

made in 2007 and then in 2011, which brought about big jumps in terms of the evaluated budget (Figure 7) (Darville, Paula; Díaz, Rodrigo; Leiva, 2017^[6]).

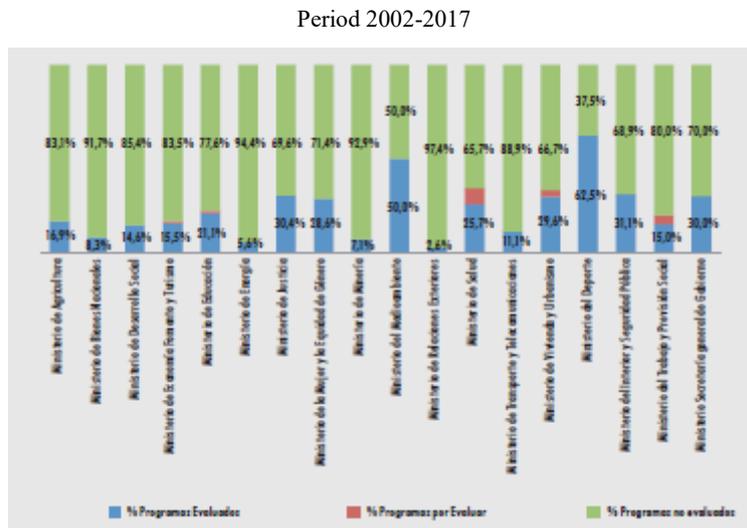
Figure 7. Coverage of Programmes and Institutions Evaluated in terms of budget



Source: (Darville, Paula; Díaz, Rodrigo; Leiva, 2017^[6]).

Finally, it is important to highlight that there is significant variation in the share of programmes evaluated across ministries. Evaluation efforts have been concentrated in the largest ministries in terms of budget and programmes (Figure 8).

Figure 8. Coverage of Programmes Evaluated by Ministry



Source: (Darville, Paula; Díaz, Rodrigo; Leiva, 2017^[6]).

The limited coverage contrasts with other OECD countries, where the unit of analysis tends to be wider. For example in the Netherlands “the MOF obliges line ministries to evaluate policy areas once every 4-7 years. Ministries are further obliged to inform the Court of Audit about evaluation analysis and conclusions (Netherlands Court of Audit, 2012). The purpose of this is to provide

better information on relevance, effectiveness, efficiency and cost of government programmes to coalition governments at the time of budget formulation. The legislation achieves comprehensive evaluation coverage of Dutch government spending.

Likewise Canada's evaluation policy has a principle of comprehensive coverage, requiring that all departments evaluate all direct programme spending within five years. Each government department or ministry has designated a head of evaluation with unencumbered access to the head of the organisation, and most have formed departmental evaluation committees. This formalisation of the evaluation function is seen to improve the neutrality of evaluations within departments. There is also a trend among a number of OECD countries towards evaluating broader policy areas. Canada Horizontal Reviews offer a good example of how to implement centrally led evaluation of a particular priority area (Box 8)

Box 8. Canada: Led performance and impact assessment of conducting horizontal reviews

The Canadian Government engages in horizontal reviews of priority areas. In support of this a Central Performance and Impact Assessment Unit (CPIAU) was created within the Treasury Board Secretariat (TBS) with the main objectives of improving horizontal measurement and delivering evidence-based insights and advice, including using statistical approaches to provide consistent and comparable measurement of programme performance. CPIAU has a strategic partnership with Statistics Canada, Canada's national statistical agency, which has a legislative mandate to collect and compile data that can strengthen programme evaluation. Statistics Canada provides information that allows datasets to be linked, thereby facilitating the horizontal approach. In addition, LMs leading the programmes being evaluated are also a part of the horizontal review team. In this sense, the Canadian experience makes it clear that, in order to have a horizontal perspective of a public policy; it is required to also have a horizontal perspective of the Government itself and to build convergence points such as the CPIAU and Statistics Canada that improve flow of information.

One of the most recent horizontal reviews was the Business Innovation Review. The review was undertaken across all departments and it aimed "to simplify programming and better align resources to improve the effectiveness of innovation programs". The Business Innovation Review was focused on a sub-set of trans-sectional Government activities that support business innovation: funding, advice, and service (existing and new); industry facing government research; support provided directly and in partnership with third parties; tech development through commercialisation and exports.

The Horizontal Review of Business Innovation and Clean Technology Programs noticed that there is a need to build a common and comparable assessment of programme performance, employing a horizontal approach, enabling greater co-ordination of policy research, evaluation approaches, and performance data, to better inform resource allocation to, and across, programmes. In addition, the Review demonstrated the potential for common measurement methods and indicators to enable comparison and benchmarking across programmes, as well as for strengthening the measurement of programme performance through statistical analysis of programme administrative data.

Source: (Treasury Board of Canada Treasury, 2018^[7])

Assessment and future orientations for reform

Chile has developed a diverse set of evaluation tools, ranging from *ex-ante* appraisal of new spending proposals to *ex-post* evaluation of individual programmes and institutions. *Ex-ante* analysis of new spending proposals has helped to improve the quality of spending. A rigorous process of analysis, involving iteration between DIPRES, MSD and LMs helps to eliminate weak proposals and improve the quality of programme design. However, there is a weak link between *ex-ante* and *ex-post* evaluation and so the potential for *ex-ante* evaluation to improve the quality of *ex-post* evaluation has not been fully realised.

The overall coverage of *ex-post* evaluations is quite small, averaging less than 5% of government expenditures by value annually. The coverage of individual evaluations is also quite small, as most “programmes” consist of small-scale policy initiatives, linked to incremental funding. This can be contrasted with other OECD countries where there are typically fewer, larger programmes being evaluated, representing major policy areas or services.

At the same time, the scope of what each evaluation covers is broad, extending beyond financial matters to consider policy and programme design, programme management and overall impact. Such a holistic approach has advantages, but it also highlights the question whether leadership of the evaluation process should be with DIPRES or another agency (see Chapter 1).

Introductory sentence:

- **Conduct evaluations of broader policy areas and groups of related programmes:** Wider evaluation coverage is likely to help identify overlaps, co-ordination challenges, and information asymmetries. It can also provide a more realistic and useful assessment of the effectiveness of programmes, since on many occasions their impact depends to a large extent in the way they interact with other associated public programmes.
- **Redefine policy criteria used to develop the DIPRES led programme of evaluations:** One suggested policy is that that all spending programmes should be subject to evaluation within a given period, for example 5 years. This change would need to be linked to reorganisation of spending programmes into fewer larger units and/or to the introduction of a new evaluation product covering sets of related programmes. A second policy change would be to consult LMs in the selection of programmes to be evaluated by DIPRES, so that the choice better reflects their concerns and priorities. This has the potential to increase LMs commitment to evaluation and the use of the results to improve programmes.
- **Develop the capacity of line ministries to conduct their own evaluations, focused on the effectiveness of policies, programmes and programme management:** DIPRES would support these evaluation efforts by LMs through developing central guidance on how to conduct evaluations. A good practice example of such guidance is the UK’s Magenta Book.
- **Create capacity within DIPRES to implement evaluations in-house.** These would be selected according to priorities determined by DIPRES and would focus on programme efficiency and effectiveness. Focused Scope Evaluations could be a useful starting point for in-house evaluations. Undertaking evaluations in-house could improve efficiency by reducing transaction costs, as well as time spent selecting evaluators, reduce implementation time, strengthen the link with the

budget sectors, and increase the use of results in budget decision-making. The UK and Ireland provide examples of how in-house evaluation capacity can be strengthened (Box 9).

Box 9. UK and Ireland. In-house evaluation capability

UK – in-house evaluation: “Evaluations are produced within departments by in-house analysts. Sometimes evaluation staff is in a segmented distinct unit, but is commonly part of existing financial management functions. Some evaluation work is tendered out to external researchers but independent evaluators outside of the government have described that administrative data can be difficult to use to evaluate the impact of government interventions.” “The government’s professional association for economists (the Government Economic Service) has also been bringing together department analysts to develop the evidence base on the trends and drivers of efficiency in the public sector in order to inform spending and implementation decisions.”

Irish – in-house evaluation specialists: “The Irish Government Economic and Evaluation Service (IGEES) is an integrated cross-government service to enhance the role of economics and value for money analysis in public policy making. Established in 2012, the Service demonstrates a desire to develop a high and consistent standard of policy evaluation and economic analysis throughout the Irish Civil Service. IGEES centralises standards of recruitment and service, to ensure that professional economic and evaluation service standards are delivered across all government departments. The IGEES Network helps build evaluation capacity in line departments. Almost all departments now have an IGEES unit with clear work programmes that should support better policy development and evaluation.”

- **Create a common methodology and unified platform to report and monitor on ex-ante evaluations:** The methodology and requirements for *ex-ante* evaluation are, in general terms, similar for both social and non-social programmes. DIPRES and MSD should reconcile their approaches to this type of evaluation. Furthermore, the two institutions should have a shared platform to report and monitor *ex-ante* evaluations. This has the potential to increase transparency, facilitate the engagement of civil society and allow for complementarities between social and non-social programmes.

3. Evaluation results and impact

The main objectives of the current evaluation system in Chile are to improve the impact of spending programmes, through review of programme design, management and financing. Impact is achieved through implementation of recommendations by the agencies responsible for the programmes and actions by DIPRES in respect of resource allocation to programmes. Chile has developed a formal process to follow-up on evaluation results and monitor implementation. However, in practice, there is low perceived impact of the evaluation system in both programme management and resource allocation. This chapter assesses the impact of evaluation results and explores the main limitations to ensure they are used as a valuable tool for decision making during the budget process formulation and for policy-making purposes.

A formal process to follow-up on evaluation results

The final stages of the evaluation process involve commitments by LMs to take action on recommendations made by evaluators, and monitoring of follow-up implementation. Within DIPRES there is also a process to determine whether there should be a budgetary response.

The recommendations made by the evaluators are studied by the Ministry of Finance and the institutions responsible for the programmes, in order to identify the way they can be implemented and the possible legal and financial limitations. Based on this analysis, institutions define formal commitments to implement the recommendations in the medium and long-term. After the evaluation process is finalised, results are shared with the Budget Director, the Management Control Division and DIPRES Budget Sectors. Evaluation reports are presented to Congress in August, before presenting the budget bill.

Institutional commitments represent the basis for the follow-up process. Some of the most common type of commitments are related to the improvement of the monitoring and evaluation system, improvement of the internal management of the programme, and substantive redesign of the programme, among others.

In order to follow-up on compliance with these commitments, DIPRES carries out a review every six months reporting on progress of institutional commitments, qualifying their degree of compliance as “fulfilled”, “partially accomplished” and “not accomplished”. There tends to be high levels of compliance with institutional commitments. More than 95% of commitments were reported as fulfilled in the last three years. Furthermore, commitments that are not accomplished or partially accomplished are usually related to a couple of programmes that are particularly problematic, rather than consistent levels of noncompliance across all programmes.

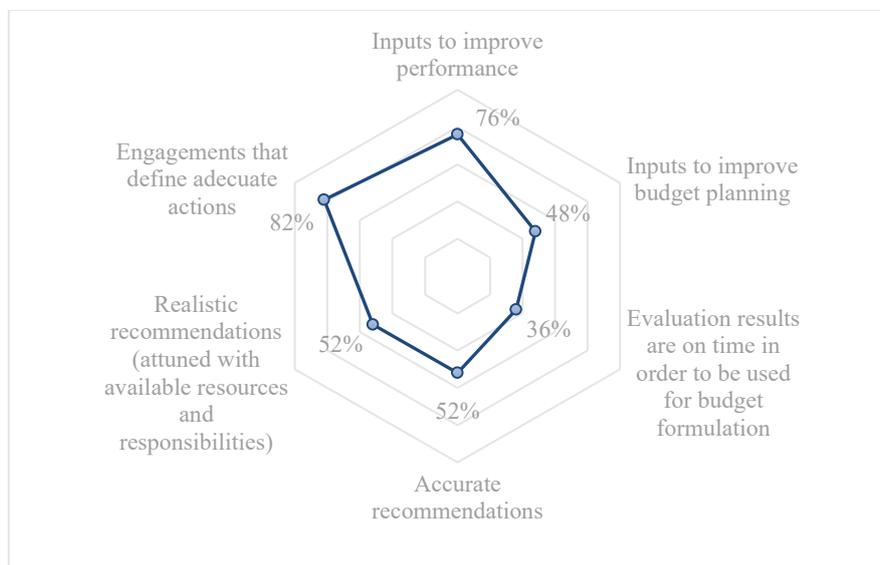
Low perceived impact on both resource allocation and policy

One of the main reasons to commission this review is the low perceived impact of the evaluation system on resource allocation and policy. This concern was echoed in discussions with stakeholders who were interviewed during the preparation of this report, representing the Presidential Administration, DIPRES, Congress, ILMs, independent evaluators, and former senior officials. Concerns were expressed that evaluations have very little impact on the existing silos or on the spending programmes themselves. Indeed, most

of the actors involved in the evaluation process agree that evaluation results are used more to define programme adjustments than to reallocate resources between or within programmes, thus the results seem to contribute more to management issues than to make the allocation of public resources more efficient, which is one of the main goals of the evaluation system.

A recent survey made by DIPRES aiming to unveil to what extent evaluation findings have an impact on resource allocation revealed that evaluation results are not largely used to inform the budget process (Figure 9). It seems that ministries are not properly absorbing and integrating the evaluation findings during the budget preparation process and the Budget Management Department of DIPRES do not find evaluation results particularly useful.

Figure 9. Characteristics of evaluation results
(Agree +strongly agree)



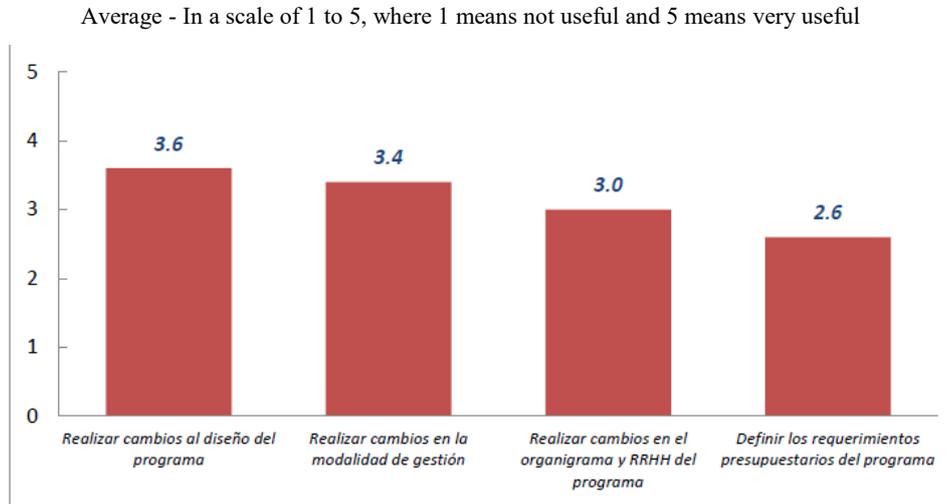
Note: Survey results from ministerial coordination officials
Source: DIPRES, 2017

Evaluation results inform processes and procedures inside programmes in order to carry out operations adjustments, and reallocate resources whether human, financial and/or material. They also inform Chilean public policies in terms of their management characteristics, the institutional arrangement and internal features of the programmes such as targeting, cost-benefit analysis, synergies and scale economies. Survey results confirm that evaluation findings are a useful tool when it comes to decision making on ongoing projects but they do not necessarily inform the agenda setting and policy formulation stages of the public policy cycle.

In a similar way, the survey also shows that evaluation results are less useful to define budgeting requirements than they are for redefining programme design, making changes in the management modality or making changes in human resources or the programme organigram (Figure 10). Regarding the way evaluations are used during the budget formulation phase, they are mostly used to justify the resources currently assigned to the programme, and to make changes in the programme’s internal processes towards improving

effectiveness and efficiency. On the contrary, evaluations are rarely used to reattribute resources between programmes.

Figure 10. Utility level of the evaluation results



Source: DIPRES

As explained before, the Chilean evaluation system includes a compliance and follow-up strategy that should inform the performance of public programmes. However, the survey highlights that the majority of respondents from the budget sector believe that the level of compliance with the evaluation recommendations (and set commitments) are not used to make decisions during the budget formulation process. Likewise, most respondents believe that there are no adequate mechanisms to inform about compliance with these recommendations. Regarding areas to improve the process, the survey highlights four main areas of reform: 1) improve the linkage between compliance and budget allocation, 2) reduce the time between the evaluation results and the definition of commitments (based on evaluation recommendations), as well as 3) improving the reporting platform to make it more flexible and interactive.

The path of follow-up actions is unclear with no process for making major cuts in expenditure and very few examples of cancellation of programmes that are performing poorly. Poorly performing programmes typically have many defenders and there are few incentives to change the programmes significantly. As underpinned by the Figure 11, more than half of the public servants that answered the survey think that the evaluations results have no demonstrated results or underperform.

Figure 11. Evaluation results by performance category

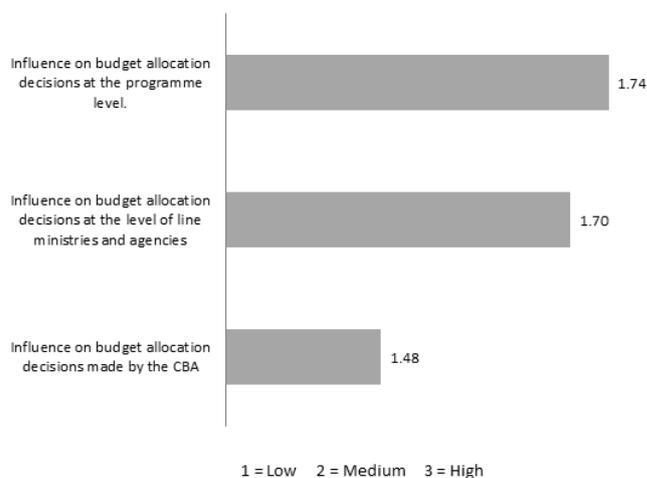
2011-2017

**Evaluation Results, by performance category
2011-2017**

Category	Total	%
Good performance	5	4%
Adequate performance	40	28%
Under performance	83	58%
No Demonstrated Results	15	10%
Total	143	100%

Source: DIPRES

A similar trend is noticed across OECD countries, where the impact of evaluation processes on budget decision making is low, especially at the macro level. Although it could be anticipated that evaluation findings would influence budget allocation decisions at the level of the Central Budget Authority, in practice this is where the influence is smaller, perhaps reflecting the decentralised nature of the evaluation process itself. Even at the programme level, where the impact is greatest, the impact level overall is rated between low and medium. The single most important problem preventing evaluations having impact is the absence of any formal mechanism to consider evaluation findings in the budget process, followed by insufficient bureaucratic and political interest in the evaluation findings (Figure 12).

Figure 12. Impact of evaluation processes on budget decision making in OECD countries

Source: OECD Performance Budgeting Survey 2018

Key challenges limiting the impact of evaluation results

Institutional arrangements

Internal co-ordination within DIPRES was flagged as an issue by several stakeholders. They noted that there is insufficient exchange of information between the budget department and the evaluation department. A better flow of information, and stronger co-ordination between the Management of Public Affairs Division and the Budget Management Department, could improve the use of evaluations results in the budget preparation process. Greater involvement by sector experts from the Budget Management Department might help to gather more and better information about the sector background and context and, accordingly, to ask better research questions.

Existing budget structure imposes a constraint on evaluations having budgetary impact

A fundamental obstacle to evaluations having budgetary impact is the current structure of the budget. The budget does not have a programmatic structure and programmes appear either as a single line item, or are hidden within a line item such as transfers. This makes it hard to identify spending allocations to programmes or to monitor and evaluate the efficiency and effectiveness of spending. While efforts have been made to prepare reports on programme execution, these have to be prepared manually, represent only a best estimate of expenditures and are not auditable. This inevitably reduces the budgetary relevance of evaluation findings. This contrasts with other OECD countries where budgets are organized and appropriated on a programmatic basis, enabling governments to track programme expenditure and link this meaningfully to performance. Moving towards a design of the budget where budgetary programmes correspond to operational programmes would allow improving the value and use of evaluations for allocation decisions.

Introducing a programme structure in the budget requires a deep cultural change, both for the CBA and for LMs. International experience suggests that programme budgeting should be implemented in a progressive manner, shifting from detailed “line-item” budgeting to programmatic and thematic budgeting, in an effort to promote greater engagement with the policy content and “meaning” of budget allocations. In many countries, programmes are selected within the context of a policy “cascade” from high strategic and developmental goals which inform medium-term and specific outcome goals, in turn informing departmental or sectoral objectives and associated output targets and deliverables. Once programmes have been selected, countries can then move to allocate clear assignment of responsibility (organisational and, ideally, managerial) for the achievement of the selected programmes and targets. Reforming the structure of the budget towards programme budgeting is an ambitious undertaking. France for example fundamentally reformed its budget structure, enacting in 2001 a new organic budget law structured along missions, programmes and actions (Box 10).

Box 10. French programme budgeting system

In 2001, France enacted a new organic budget law including a well-defined programme structure, shifting budget classification from nature of expenses to public policy objectives. According to this new approach, the budget must be divided into missions, programmes and actions:

- A mission covers a series of programmes designed to contribute to a specific public policy. A mission can involve a single ministry or several ministries. The Parliament cannot change or adjust the Missions. It has to accept the budget allocations proposed by the executive government and has power only to vary the allocation between programmes.
- A programme covers a coherent set of activities of a single ministry targeted to a specific public policy objective. If more than one ministry participates in a large public policy, each of them should have a separate Programme, covering its own responsibility in that matter, and ensuring coordination. Thus a programme corresponds to a centre of responsibility. Accordingly, for every programme, a programme director is appointed. All the resources from the State Budget should be allocated and spent within a Programme. In a similar way, resources allocated by the Parliament to a particular programme cannot be spent by the ministers for another programme.
- An action covers a set of operational means to implement the programme. The budget breaks down resources allocated to the actions of each programme; however, this break down is indicative and not committing. There is indeed a high degree of freedom for expenditure choices for ministers, in order to allow the Programme to reach its forecasted performance. However, there is one exception to this increased freedom: appropriations for personnel are not indicative but binding, in an asymmetrical way: personnel appropriations can be used for other purposes, but appropriations for other purposes cannot be used for personnel costs.

The Organic Budget Law prescribes an extensive performance reporting process to integrate performance information in the budget system through the following two types of mandatory budget documents: annual performance plans (*projets annuels de performances*, PAP) and annual performance reports (*rapports annuels de performances*, RAP). For a given mission, the PAP provides a detailed description of its purpose, goals, policy targets and performance indicators. As part of the annual budget act, the PAP documents are forward looking and are meant to contribute to the public debate about the costs and benefits of public policy. The RAPs are published in the first quarter along with the budget review act; they focus on performance achievements and provide detailed information on programme implementation and results. The RAPs are thus backward looking and tend to contribute to the public debate on the administration's performance.

Source : Loi organique relative aux lois de finances 2001

Loss of credibility on the evaluations system and relevance of results

As already mentioned, according to the survey there is a generalised perception that LMs do not pay enough attention to the evaluation results. There seems to be a lost momentum for the evaluation system in Chile. Given DIPRES's strong position in defining policy making and resource allocation within the Chilean budgeting framework, DIPRES role as the leader of the evaluation system should be perceived as structural. This was the case in the first years of the evaluation system. However, with the passing of time, results and findings were perceived as having low impact in decision making. As a consequence, the level of attention paid to the evaluation by line ministries became weaker.

Lack of specific tools to increase efficiency and improve budget allocation

In contrast to other OECD countries, Chile has invested heavily in the evaluation system but has not developed other performance monitoring and management tools that typically form part of an overall performance management and budgeting system. Some rebalancing and broadening of the effort seems likely to produce better results. The list of tools typically used by OECD countries includes:

- reforms of ministry strategic plans to improve the quality of programme objectives, performance indicators and targets,
- use of a medium-term expenditure framework, to align sectoral spending plans with policy priorities consistent with meeting medium term fiscal objectives,
- spending reviews, to identify inefficiencies in existing spending and free up resources for new priorities,
- programme and performance budgeting, defining programmes and performance indicators and linking these to resource allocation, and changing budget classification and expenditure controls
- reform of HR management to align individual incentives with performance, for example performance-based pay and individual performance appraisal.

Chile has used some of these tools, including a strategic definition of institutions functions, objectives and performance indicators (*definiciones estratégicas*), a management improvement programme (*Programa de Mejoramiento de Gestión*), and salary incentives (*Mecanismos de Incentivo de Remuneraciones*). However, there is other tools that could be further developed. Given the current political and institutional context two of these tools; programme budgeting and spending reviews would appear to have strong potential to complement and strengthen programme evaluation.

Programme and performance budgeting

Performance budgeting is “the systematic use of performance information to inform budget decisions, either as a direct input to budget allocation decisions or as contextual information to inform budget planning, and to instil greater transparency and accountability throughout the budget process, by providing information to legislators and the public on the purposes of spending and the results achieved”

Performance budgeting offers a wide range of potential benefits that are attractive to different stakeholders. For the centre of government, performance budgeting offers the chance to strengthen the alignment between decisions on budget allocation and government

policy priorities, thereby boosting the chances that the government can deliver on important pledges. For parliaments, performance budgeting offers greater clarity on the purposes of spending and what goods and services will be delivered in return for the resources they have voted on, as well as a means of holding officials to account for the achievement of results. For Ministries of Finance, performance budgeting provides new types of information that help them make resource allocation decisions based on evidence of what works, plus tools to make line ministries more accountable for the effectiveness and efficiency of spending. For line ministries, performance budgeting provides tools to improve their internal decision making, and to make a stronger case to government in support of their budget proposals. It also helps programme managers to do their job, enabling them to track performance as well as spending. For citizens and civil society organisations, performance budgeting offers the prospect of greater transparency and accountability in respect of the purposes and results of public spending and, by opening up the ‘black box’ of the budget, it strengthens the basis for direct citizen engagement in the budget process.

Characteristic elements of a performance budgeting system are the following:

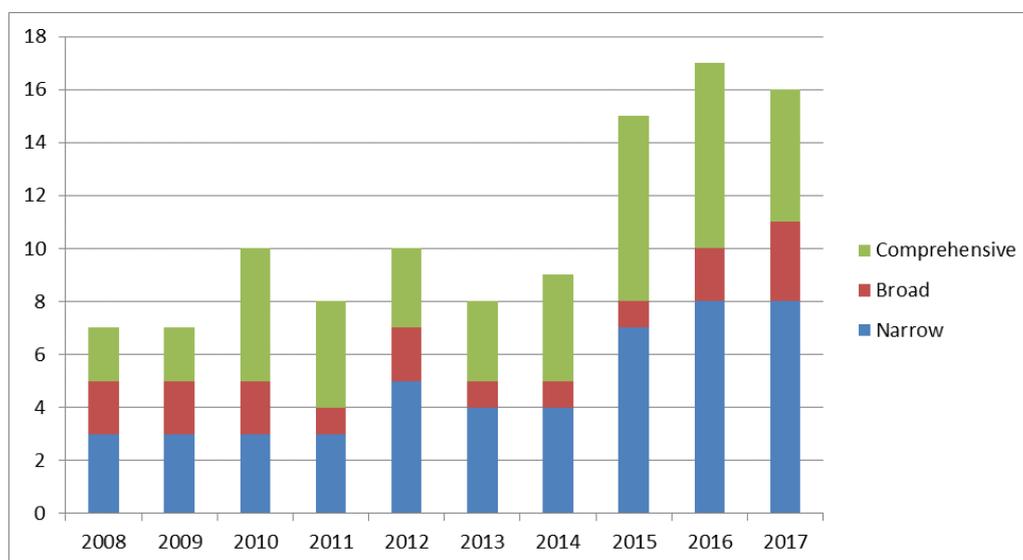
- In the preparation of the budget, spending decisions take account of priority policy objectives as well as past programme performance;
- The budget is presented to the legislature as a set of programmes. This may be in addition to, or replace, presentation based on administrative and economic classification;
- Non-financial performance indicators and targets, representing the planned objectives of spending, are presented in the budget and linked to spending programmes;
- During budget execution, budget managers have the autonomy to manage financial resources, balanced by the accountability for achieving results;
- Performance information is provided to managers together with budget execution data to help them monitor performance and improve the quality of spending;
- Reports to parliament on budget execution include performance information and a narrative explaining the reasons for under or over-performance.

Performance budgeting is a major long-term reform which deserves separate analysis and recommendations. In the short term, implementation of a programmatic classification of the budget, which is one element of the reform, would substantially strengthen the evaluation process, independent of full implementation of a performance budgeting system.

Other elements of the reform would strengthen performance orientation and complement the evaluation process. DIPRES has already begun work to define programmes.

Spending Reviews

Spending reviews provide a method of reviewing spending effectiveness and efficiency linked to government policy goals. They are widely used as a strategic budgeting tool in OECD countries with all but 5 countries reporting that they conduct spending reviews either annually or periodically. Between 2011 and 2018 there has been a steady trend towards the use of spending reviews in OECD countries (Figure 13).

Figure 13. General trends towards increased use of all types of spending reviews

Source: 2018 OECD Performance Budgeting Survey

The main objective of spending reviews is to increase the fiscal space available to government to finance its policy priorities. Governments have made use of this tool to carry out budget consolidation, cutting less effective spending while safeguarding key spending priorities and to reprioritise expenditure, helping governments to align spending with key policy priorities. A key characteristic of spending reviews is that they systematically analyse baseline expenditures to identify scope for savings and realignment of resources to support policy priorities. This contrasts with the normal focus in the annual budget process on competing demands for incremental increases in spending.

Spending reviews have many similar characteristics to evaluation, including a critical re-assessment of the justification for expenditures, from the perspectives of economy, efficiency and effectiveness. Spending reviews are an adaptable tool which can be tailored to the specific circumstances of individual countries and may examine specific programmes, policy areas, or agencies. The use of narrow spending reviews has increased constantly in OECD countries since 2008 (Figure 13).

Spending reviews also have the potential to increase the relevance and impact of evaluations by making use of their findings to inform budget decision making. In Canada “strategic reviews” have raised the profile of the evaluation function by requiring departments to systematically address fundamental issues of programme relevance. In internal reviews of evaluation, two-thirds of programme managers reported that evaluations were useful or somewhat useful for spending reviews (Government of Canada Treasury Board Secretariat, 2013).

Smaller-scale sector spending reviews, such as those used by Denmark and the Netherlands offer a model that could be followed by Chile (Box 11).

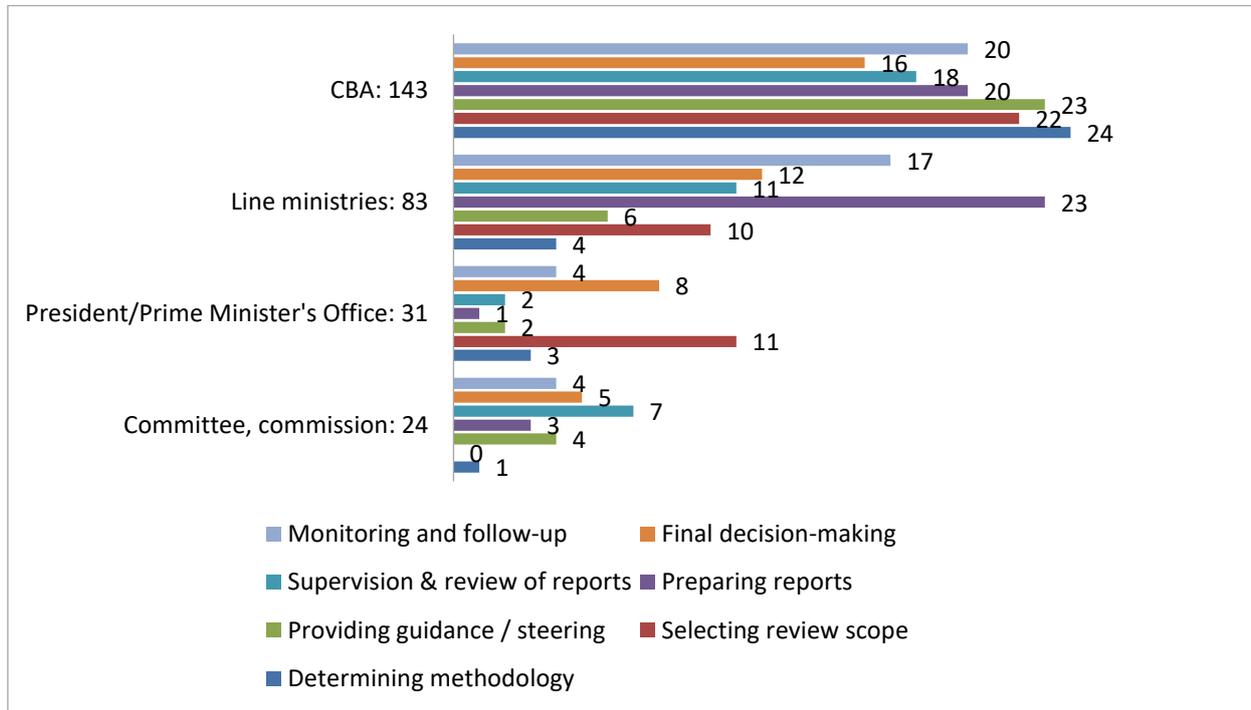
Box 11. Denmark: Special Studies review

Since the mid-1980s, Denmark has had a system of spending reviews known as “special studies”, and this process has continued to operate right up to the present time. The special studies are part of the normal annual budget preparation process, although there have been some years when no special studies have been undertaken. Although in principle the special studies may recommend increases in funding for existing programmes, in practice this is rare and the focus is upon savings measures. There is no formal link to broader government performance-improvement processes. There have typically been 10-15 special studies carried out each year, although this has increased significantly since the GFC. Historically, the primary focus of the special studies process has been upon increasing space for new expenditure priorities. However, at the present time the focus has shifted more towards aggregate expenditure reduction for fiscal consolidation purposes. This has led to an increase in both in the number of special studies and in the value of expenditure which they cover (e.g. studies of defence and police expenditure). Most special studies are agency reviews or programme reviews, and the main focus has, over the years, been upon efficiency savings rather than strategic (output) savings. Special studies are generally carried out by joint MOF/spending ministry taskforces, with formal terms of references approved by Cabinet. Taskforces present savings options to the Minister of Finance and the Economic Committee of Cabinet. These recommendations should in principle be based on consensus between the MOF and the spending ministry concerned, but if consensus is not reached, separate recommendations may be put forward. The Economic Committee generally makes the final decision about which savings measures will be adopted in the budget.

Source: Spending reviews, OECD Journal on Budgeting Vol 2013/2

CBAAs play a leading role at almost all stages of the spending review, especially in determining the methodology, while LMs lead the preparation of the reports and the President or PM office are key actors when selecting the review scope (Figure 14).

Figure 14. Primarily responsible government actor for spending review procedures (2018)

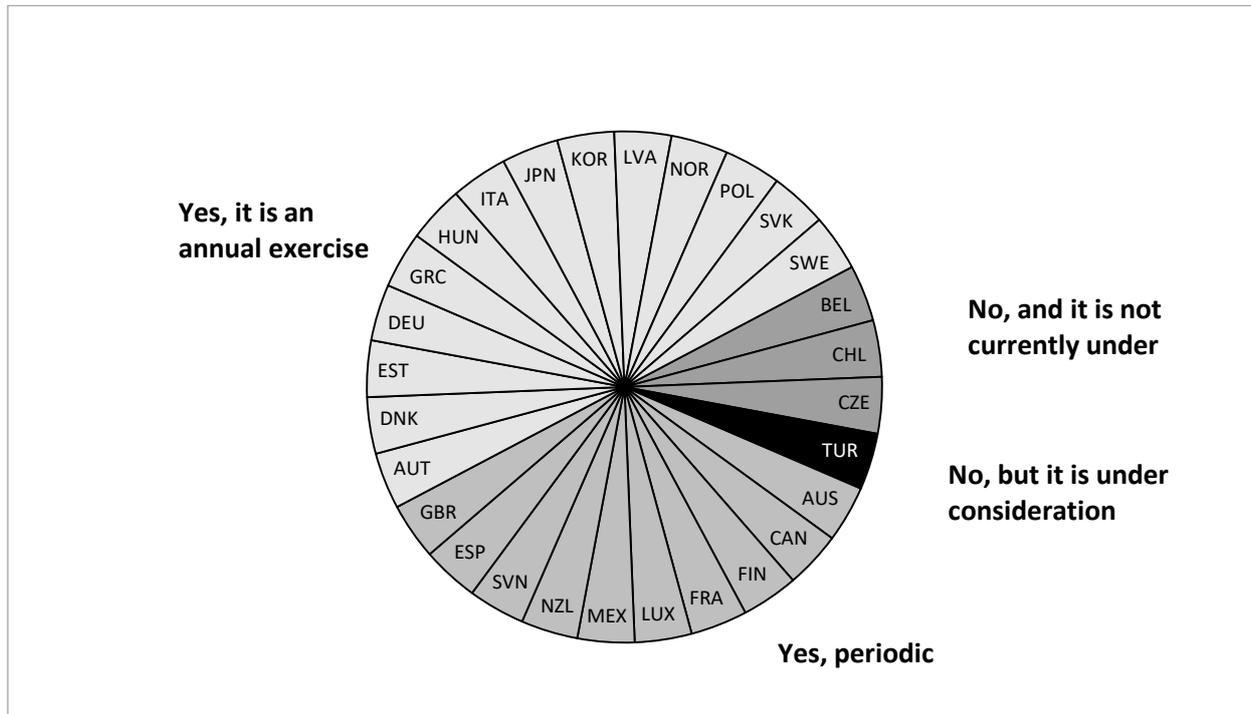


Source: 2016 OECD Performance Budgeting Survey

Finally, in terms of information available, it is important to consider what information exists already in terms of outcome evaluation (using existing relevant performance indicators) and/or efficiency analysis. For this, Australia provides a positive example through their system of “Strategic Reviews” of programmes or processes that have been conceived in part to support spending review. Drawing on these reviews, Department of Finance typically develops and presents savings options for the Expenditure Review Committee of Cabinet.

It has been a widespread use of spending review in OECD countries. In 2018, half of OECD countries implement it as an annual exercise and 10 countries use it as a periodic exercise (Figure 15)

Figure 15. Application of spending reviews (2018)



Source: 2018 OECD Performance Budgeting Survey

Assessment and future orientations for reform

Strengthen political support for evaluation

Government leadership is key to developing a culture of accountability for performance would give renewed energy to the evaluation process. In other OECD countries political expression has taken different forms. The United States and Australia, for example, have enacted performance laws (Australia PGPA Act 2103, US Government Performance and Results Modernization GPRMA Act 2010), whereas countries such as the UK and Austria set up Performance Units at the centre of government. Political engagement and communication could be linked to decisions regarding changes to the evaluation system.

Implement a programmatic budget classification

The impact of evaluations on the budget is affected by misalignment between programmes, which are the main unit of evaluation, and the existing budget classification system, and the businesses and products of the line ministries. Transitioning from the current budget structure, based on administrative classification, towards a programmatic structure and combine existing budget programmes into larger programmes. Existing programmes could be reclassified as sub-programmes or activities. This would improve the basis for evaluation by enabling programme expenditures to be tracked. Programmes should cover broad policy areas and services that have a long-term character, for example, in education programmes might include primary education, secondary education, adult literacy, special needs education, teacher training etc. DIPRES has already begun work to define programmes, and this should be continued in collaboration with line ministries but with a clear government commitment and timetable to adopt a programmatic classification.

Introduce spending reviews as a complementary tool to evaluation

DIPRES should consider initiating spending reviews modelled on those used by Denmark, the Netherlands or Canada. These would focus on groups of programmes or policy areas and would be focused on programme efficiency and effectiveness and opportunities to improve strategic alignment of spending with priority policy goals.

As highlighted in the Budgeting in Chile report (Vammalle and Ruiz Rivadeneira, 2017^[8]), OECD best practices suggest the following key elements for implementing a successful spending review process in Chile:

- 1) Spending review should be a regular an ongoing feature of the budget process. Simultaneous consideration makes it possible for governments to adopt other high priority new spending proposals without increasing aggregate expenditure, by selecting additional savings options sufficient to fund the extra new spending.
- 2) Clear definition of the scope of the spending review. A proper alignment between the objective of the government and the type of review implemented is a key success factor for the spending review exercise.
- 3) Political oversight and direction of the spending review process. Spending review is not a purely technical function; it should be under the direction of politically-appointed officials who are sensitive to the priorities of the Presidency.
- 4) Clear definition of roles in the spending review process. Taking into account the clearly centralised evaluation process currently in place, Chile could benefit from a joint spending review exercise, where line ministers can contribute with their particular knowledge of the sector and DIPRES can support and orient the process due to its long tradition with the evaluation and control system.
- 5) Ensure line ministries buy in. Increase efforts to promote the evaluation function within spending institutions could be a good strategy to increase engagement and support from line ministers.
- 8) Have a well-defined programme budgeting. A successful spending review depends in large extent on the MoF knowing exactly which programmes are being delivered by each spending ministry and how much is being spent on each, making the programme classification of the budget particularly relevant to program review, and to the search for strategic savings (Robinson, 2014^[9]).
- 9) Implement the review at the right moment in the political cycle and following a clear implementation timeline (Box 12). Traditionally countries carry out in-depth spending reviews in the first year of the presidential term of office, and more limited spending reviews in subsequent years. It is helpful to agree the sequencing of spending reviews, for example beginning with “quick wins’ that may demonstrate the value of the exercise and so generate momentum and commitment to the process.

Box 12. Spending Review process and main steps

Phase 1 - Mandate and Road Map

Secure a political mandate and bureaucratic commitment to the exercise supported by communicating objectives and assigning leadership roles. At the outset it is important to establish how decisions will be made based on the findings of the spending review process. Significant changes in resource allocation will undoubtedly meet with resistance and so high level support will be needed to realise the potential benefits.

Develop a high-level roadmap or concept note encapsulating agreement on the scope, objectives and timescale for the spending review exercise. The objectives should reflect national policy priorities and fiscal goals.

Allocate resources dedicated to the project. Senior officials will have to be released from their usual responsibilities. High level analytical and policy skills are needed to assess potential reform options and top-level project management skills are needed to keep the process on track and to ensure consistency. Operational leadership of the project should remain with government and not outsourced. Knowledge of administrative rules and processes is necessary but not sufficient and external experts may need to be hired.

Phase 2 - Design

In Phase 2 the mandate and road map are operationalised. High-level objectives and scope are disaggregated into a portfolio of smaller projects and more detailed terms of reference are prepared for each component part. In case the Government has overall savings targets, the expected contribution of each line ministry should be communicated up front. Line ministries may be invited to suggest spending items or agencies that should be subject to particular scrutiny.

The key output should be a set of terms of reference that combine a unified overall approach or methodology with more detailed sector-specific requirements. Methodological guidance is advisable to ensure a consistent approach across the portfolio of individual spending reviews, including definition of the technical elements required to support decision making, such as potential savings, costs, impact on citizens, risk assessment etc.

Governance arrangements should be defined and a team assembled for the overall co-ordination of the spending review exercise before the implementation phase begins. A co-ordination/delivery taskforce is typically a permanently staffed team reporting to a board which has a consolidated view over the entire project during the conduct of the review and the implementation phase. Independence from the administration under review is a key success factor, so that decisions are not made by officials responsible for the policy or spending under review.

Phase 3 – Conduct of the Reviews

Conduct of the reviews incorporate three key steps diagnosis, formulation of reform options, feasibility analysis and decision making. Diagnosis is based on mapping of sources of funding and expenses including analysis of spending evolution and the main cost drivers, broken down to a level where potential inefficiencies and anomalies can be identified. Spending should be assessed in the context of policy objectives and its relative contribution towards achieving policy objectives including target populations. Useful perspectives can be obtained through benchmarking against national or international comparators.

Lack of data is a common problem in which case alternative approaches can be explored and combined. These may include use of sample data and other existing material e.g. from auditors, universities think tanks etc. A good analysis will challenge historically-accepted assumptions and procedures that are obsolete.

The second key step is the formulation of reform options with alternative scenarios. This may involve sensitivity analysis to simulate the impact of a lower level of spending and/or of an alternative allocation of spending on policy objectives. Another approach is to consider policy alternatives. Some countries consult widely with officials and end-users to generate ideas or voice their preferences between different proposals.

The third key step is to analyse the feasibility of the various reform options. Feasibility and impact are the main criteria used to decide between alternative options. Feasibility analysis should take account of multiple factors including implementation cost, risks (political and public resistance), other obstacles such as the need to amend the law and spill-over effects e.g. on other expenditures etc.

At the end of this phase, the recommendations of each individual spending review are considered and a set of reform proposals is selected for implementation based on considerations of impact, cost, risk and return.

Phase 4 - Implementation

In this phase the benefits of the spending review are realised in the form of savings and reallocation of resources. Follow-up monitoring is as an essential part of the process as it should not be assumed that recommendations will be implemented fully and effectively. Results should be assessed both individually for each project and collectively relative to the initial mandate.

At the end of the implementation phase it is recommended to evaluate the real savings captured versus the implementation cost and the fulfilment of the initial mandate. This exercise should be carried out independently. For example the UK National Audit Office independently assessed the savings achieved by the 2004 Gershon spending review and the Finnish National Audit Office (VTV) published an evaluation of the Productivity Programme.

Source: Adapted from (Vandierendonck, 2014^[10])

Strengthen internal coordination within DIPRES

The feedback loop, whereby evaluation findings influence subsequent budget decisions is weak. In part this is due to inherent problems in linking budget allocations to outputs and outcomes, caused by the complexity of the problems that governments deal with and the often unclear relationships between financial inputs, outputs and outcomes. In part however it can be attributed to poor internal co-ordination between the departments in DIPRES that deal with evaluation and budget, and limited follow up with LMs.

DIPRES management should institute formal mechanisms to ensure effective information exchange between departments within DIPRES to improve the quality and impact of evaluation. In the course of designing and conducting evaluation the knowledge and concerns of sector budget experts should be considered. Equally the findings of evaluations should be considered by sector specialists in the budget department responsible for preparing the next year's budget.

4. Summary findings and recommendations

The system of *ex-ante* and *ex-post* evaluations developed by the DIPRES has many good features that reflect the Congress and government's strong and sustained commitment to evaluate spending programmes and feed the results back into the budget decision-making process. Strong features include rigorous *ex-ante* analysis of programmes and well developed methodological approaches to *ex-post* evaluation. At the same time the system has not produced the expected benefits, in terms of either improved accountability for programme performance or changes in expenditure patterns and so some re-thinking is needed. Several broad set of issues that adversely affect the system have been identified in the preceding analysis.

First, Chile has invested heavily in the evaluation system as the principal tool for performance monitoring. It stands in contrast to most other OECD countries that use a more diverse set of performance monitoring and management tools to track programme performance and make budgetary adjustments based on performance. Some rebalancing of the tools used by DIPRES would likely produce better results. This would involve:

- Development of performance budgeting,
- Selective use of spending reviews
- Interim performance monitoring, especially of major infrastructure programmes

Second, the evaluation system is highly centralised, with limited involvement of line ministries, Congress and other stakeholders in decisions about what should be evaluated and in the conduct of evaluations. This creates problems of ownership of evaluation findings and the relevance and quality of the analysis. A broadening of involvement in evaluation, to include other stakeholders, could lead to improved quality and impact. More specifically this would involve:

- DIPRES leading evaluations of important strategic programmes, in consultation with LMs and Congress.
- Line Ministries developing the capability to conduct their own evaluations, following central guidance provided by DIPRES, and to participate in DIPRES led evaluations.
- The CGR developing a capability to carry out performance audits, responding especially to requests from Congress

Third, mechanical problems in the design of the budget system, especially the fact that programmes are not well defined in the budget, make it hard to track programme expenditure and therefore hard to draw budgetary conclusions based on evaluation findings.

- Introduce a programme classification in the budget, transitioning from the current structure, based on administrative classification.
- Combine existing budget programmes into fewer, broader policy areas designating smaller existing programmes as sub-programmes or activities. This is likely to help identify overlaps, co-ordination challenges, and information asymmetries

Fourth, there has been a long-running debate regarding oversight and independence of the evaluation function. Congress, while having a strong legal role, has felt relatively powerless in practice, while civil society has also had a limited role, despite evaluation reports being publicly available. Rather than establishing a separate independent evaluation function this review recommends measures to strengthen independent oversight of the evaluation process:

- Strengthen the analytical capacity of the Congress (secretariat), to support stronger engagement of members at all stages of the evaluation process.
- Establish a Civil Society Council to independently review and comment on budget matters including evaluation.

There is a general concern about the lack of impact that evaluations have on budgetary decision making, or on decisions to redesign or cancel programmes. This relates to co-ordination issues within DIPRES and between DIPRES and MSD. OECD recommends to:

- Establish formal internal co-ordination mechanisms within DIPRES, between the departments responsible for evaluation and budget preparation. Budget analysts should be involved in scoping evaluations led by DIPRES and consider evaluation findings as part of the process for setting the next year's budget for programmes subject to review.
- Create a common methodology and unified platform to report and monitor on *ex-ante* evaluations.
- Link *ex-ante* evaluation more firmly to *ex-post* through the identification of objectives, indicators, an overall logical framework and a tentative timeframe for future evaluations.

Finally, at a more fundamental level successive government have shown limited interest in programme performance, with the result that the main actors lack incentives to pay attention to evaluation findings. A renewed political commitment and messaging, as part of a general overhaul and re-launch of the evaluation system, would help improve the impact of future investment in programme evaluation.

References

- Castro, M., G. Lopez-Acevedo and G. Busjeet (2009), *El Sistema de M&E de México: Un salto del nivel sectorial al nacional*, <http://www.worldbank.org/ieg/eed> (accessed on 11 December 2018). [3]
- Darville, Paula; Díaz, Rodrigo; Leiva, J. (2017), *Cobertura evaluación de programas e instituciones públicas*, Dirección de Presupuestos, Ministerio de Hacienda. [6]
- Irarrázaval, I. and B. De Los Ríos (2014), “Monitoreo y evaluación de políticas públicas Contribuciones de la experiencia internacional para la institucionalidad chilena”, *NOTAS PÚBLICAS 2*, <http://www.politicaspUBLICAS.uc.cl>. [4]
- Irarrázaval, I. and B. De Los Ríos (n.d.), *NOTAS PÚBLICAS Monitoreo y evaluación de políticas públicas Contribuciones de la experiencia internacional para la institucionalidad chilena*, <http://www.politicaspUBLICAS.uc.cl>. [12]
- OECD (2018), *Gender Equality in Canada: Mainstreaming, Governance and Budgeting*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264301108-en>. [2]
- OECD (2014), *Chile’s supreme audit institution: enhancing strategic agility and public trust.*, OECD Public Governance Reviews. [1]
- OECD (Forthcoming), “Budgeting Practices and Procedures in OECD countries”. [5]
- Robinson, M. (2014), “Spending reviews”, *OECD Journal on Budgeting*, Vol. 13/2, pp. 81-122. [9]
- The World Bank (2009), “Making Smart Policy: Using Impact Evaluation for Policy Making. Case Studies on Evaluations that Influenced Policy”, *Doing Impact Evaluation* 14. [11]
- Treasury Board of Canada Secretariat (2018), *Centrally-led Performance and Impact Assessment*. [13]
- Treasury Board of Canada Treasury (2018), *Centrally-led Performance and Impact Assessment*. [7]
- Vammalle, C. and A. Ruiz Rivadeneira (2017), “Budgeting in Chile”, *OECD Journal on Budgeting*, Vol. 16 (3), pp. 9-70. [8]
- Vandierendonck, C. (2014), “Public Spending Reviews: design, conduct, implementation”, *EUROPEAN ECONOMY Economic Papers* 525. [10]