OECD Public Governance Policy Papers

Improving decision making through policy evaluation in Portugal

OECD
BETTER POLICIES FOR BETTER LIVES

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

This work was approved and declassified by the Public Governance Committee and the Committee of Senior Budget Officials on 31 October 2024.

Photo credits: Cover © Xinovap/Shutterstock.com.

© OECD 2024

Attribution 4.0 International (CC BY 4.0)

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (https://creativecommons.org/licenses/by/4.0/).

Attribution – you must cite the work.

Translations – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of original work should be considered valid.*

Adaptations – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

Third-party material – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

Acknowledgements

The paper was written by Claire Salama, Monitoring and policy evaluation lead in the Public Management and Budgeting Division of the Public Governance Directorate and Silvia Picalarga, Junior Policy Analyst. This work was conducted under the leadership of Jón R. Blöndal, Head of the Public Management and Budgeting Division, and the overall steering of Elsa Pilichowski, Director of the Public Governance Directorate. Meral Gedik prepared the paper for publication.

This paper is part of a project conducted in 2022 by the OECD with the Competence Centre for Planning, Policy and Foresight in Public Administration in Portugal (PlanAPP). This project was co-ordinated by Paulo Magina, Head Infrastructure and Public Procurement Division (IPP) and would not have been possible without the support of PlanAPP. The OECD Secretariat especially thanks Paulo Areosa Feio, Director, Maria Oliveira Fernandes, Policy Advisor, and Catarina Cristino Pereira, Head of Unit of Policy Evaluation as well as her team, for the support and close collaboration throughout the whole policy review process.

Table of contents

Acknowledgements 3

Executive summary 5

1 Institutionalising policy evaluation from a whole-of-government perspective 7
- 1.1. Creating a shared understanding of policy evaluation 8
- 1.2. Cross-government high-level guidance on policy evaluation 10
- 1.3. Attributing institutional mandates inside the executive 12
- 1.4. Promoting a rich evaluation system outside the executive 14

2 Promoting quality of policy evaluations 16
- 2.1. Engaging with evaluation stakeholders and users 16
- 2.2. Fostering capacities and skills for policy evaluation 17
- 2.3. Availability of, and access to, quality data 19
- 2.4. Ensuring good governance and technical quality of evaluations 22

3 Conducting policy evaluations that impact decision making 25
- 3.1. Embedding evaluations in decision-making processes 25
- 3.2. Publicity and communication 27

4 What role for an evaluation champion? 30
- Summary of Recommendations 33

References 35

FIGURES

Figure 1.1. Definition of policy evaluation 9
Figure 1.2. Institutions within the executive that have competences related to policy evaluation across government 13
Figure 1.3. Actors outside the executive with role in policy evaluation 14
Figure 2.1. Mechanisms used for the promotion of quality in OECD countries 22
Figure 3.1. Publicity of policy evaluations 28

Executive summary

Using evidence for decision making is crucial to achieving societal and economic goals, ensuring governments' resilience and preserving trust in public institutions. As countries are confronted with increasingly complex challenges – such as climate change, shrinking fiscal space and multiple crises – evidence on what works is needed to guide government action. Evidence-informed decision making contributes to government agility and resilience in uncertain contexts.

In recent years, Portugal has strengthened its ability to take evidence-informed decisions by establishing the Portuguese Competence Centre for Planning, Policy and Foresight, or PlanAPP. This agency was created to help the Portuguese public administration deliver better policies and brings together different actors of the administration, academia and civil society. In particular, the *Unidade Técnica de Avaliação* (UTA) or Evaluation Unit of PlanAPP plays a co-ordination role in the Portuguese evaluation system by promoting the evaluation of public policies; contributing to the systematisation, development and dissemination of guidelines on the evaluation of public policies; ensuring the process of evaluating and quantifying the economic, social and environmental impact of public policies; and promoting collaboration with relevant external entities. Yet, at the time of writing this paper, the Portuguese government still suffers from challenges that hinders an evidence-informed decision-making approach.

Sound institutional set-ups help ensure that evaluations are systematically conducted. In Portugal, there is little institutionalisation and systematisation of evaluations across the government, and, as a result, little opportunity for actors to share good practices across the system, to conduct cross-sectorial evaluations and to create a critical mass of skills. To improve the systematisation of policy evaluation, the paper recommends that PlanAPP develop high-level guidance in the form of a multi annual cross-government evaluation agenda in consultation with line ministries. In addition, PlanAPP would benefit from the development of a legal framework providing clear mandates for policy evaluation, defining periodicity of evaluations and their criteria.

The quality of evaluations is important to ensure that they are used in decision making and can support learning, accountability and better decision making. Currently, the quality of evaluations conducted in the Portuguese public administration could be improved, because line ministries generally lack the necessary skills, data and quality mechanisms to produce high-quality evaluations. Indeed, investing in training and building capacity for conducting and commissioning policy evaluation will be essential if PlanAPP is to improve the quality of its evaluations.

In addition, access to quality data plays a fundamental role in evaluation. There is need to improve the governance of data to allow both to PlanAPP and line ministries to access microdata for evaluative purposes.

To further strengthen the quality of evaluations, a network of government evaluators could be established to share good practices and methods on policy evaluation together with a peer review group to support line ministries in ensuring the quality of their evaluations. PlanAPP could also further cultivate networks with academia to both promote the supply of evidence for policymaking and improve the impact of scientific and research evidence.

Finally, ensuring that evaluation results affect decision making is an essential element of an effective evaluation system. The impact of evaluation results is currently difficult to assess, as evaluations are not always communicated or linked to decision-making processes. To improve the impact of evaluations it will be important to ensure that policy evaluations are made public by default, as is already the case for evaluations of EU-funded programmes and projects. To ensure that evaluations are not only available but are easy to find, PlanAPP could develop a user-friendly evaluation database to facilitate access to results. In addition, PlanAPP could support line ministries in better linking evaluations to other decision-making processes to strengthen the impact of evaluation results.

Overall, there is an important opportunity for PlanAPP to play a central role in strengthening the Portuguese public policy evaluation system by fostering a co-ordinated approach across government and promoting demand for evaluations and evidence from decision makers.

Building on the OECD Recommendation on Public Policy Evaluation (see Box 1.1), this paper analyses the current Portuguese evaluation system and provides tailored recommendations on how to strengthen the institutionalisation, quality and impact of policy evaluations across government, with a particular focus on PlanAPP's role in contributing to these objectives. In doing so, it relies on a gap analysis of Portugal's current practices compared to the standards that are set out in the OECD Recommendation on Public Policy Evaluation. It also draws on good practice examples from other member countries and those of France, Lithuania, the Netherlands and the United Kingdom in particular. As such, the paper is structured following the key principles set out in the Recommendation: the institutionalisation of policy evaluations from a whole-of-government perspective, practices that promote the quality of policy evaluations, and the extent to which Portugal conducts evaluations that impact decision making. Finally, the last section of the paper offers perspectives on the role that PlanAPP could play in championing policy evaluation practices across the system.

1 Institutionalising policy evaluation from a whole-of-government perspective

Institutionalising policy evaluation from a whole-of-government perspective is essential to ensure that isolated and unplanned evaluation efforts can become more formal and systematic approaches. There is no perfectly uniform approach to establishing a government-wide institutional framework for policy evaluation. However, the OECD Recommendation on Public Policy Evaluation (see Box 1.1) underlines that countries need to undertake three key actions to make sure that policy evaluations are conducted systematically and with a whole-of-government perspective (OECD, 2022[1]):

- Define and **assign institutional responsibilities** for conducting policy evaluations.
- Define **high-level guidance across government**, such as in a legal or policy framework or in a multi-annual evaluation agenda.
- Foster a culture of learning and accountability by **promoting demand** for, and ownership of, evaluations within and **beyond the executive**.

Box 1.1. The OECD Recommendation on Public Policy Evaluation

The OECD Recommendation on Public Policy Evaluation was adopted by the Council in June 2022 (OECD, 2022[1]). The Recommendation represents the first international standard aimed at driving the establishment of robust institutions and practices that promote the use of public policy evaluations. It is structured around three mutually reinforcing pillars: institutionalisation, quality and impact:

- I**nstitutionalise public policy evaluation from a whole of government perspective by:**
  - Conduct and use evaluations across government ensuring that they are carried-out in a systematic manner and that their results are used in policy and budgetary decision making.
  - Foster a culture of learning and accountability by promoting demand for, and ownership of, evaluations within and beyond the executive.
- **Promote the quality of public policy evaluations by:**
  - Actively plan, design and manage evaluations so that they are timely and proportionate to the intended objectives, taking into account the needs of the primary users and the types of intended uses, and ensuring that results can be trusted by stakeholders.
  - Establish quality standards and mechanisms for evaluations to generate robust and credible evaluation results that can be trusted and used with confidence
  - Develop institutional skills and capacities to conduct, commission and use evaluations effectively and in a credible manner

- **Conduct public policy evaluations that impact decision making by:**
  - Establish institutional mechanisms to embed evaluation in decision-making processes, both at the organisational level and across government.
  - Provide easy access to evaluations and present the findings deliberately in order to improve the uptake of evaluation results.

Source: (OECD, 2022[1])

Together with these three elements, empirical evidence suggests that having of a clear and shared definition on what policy evaluation is can help improve the institutionalisation of the practice by creating a common understanding across government (OECD, 2020[2]).

In the Portuguese government, several actors both within and outside of the executive have mandates to conduct evaluations. This is the case of the Portuguese Court of Auditors, as well as the Development and Cohesion Agency (AD&C). In addition, the establishment of PlanAPP, which has an explicit mandate for co-ordinating policy evaluations across government as set out by the decree n.21/2021, represents an important step in promoting a cross-government approach to this practice. Yet, several elements hinder a more systematic approach to evaluations across-government and the creation of a sound policy evaluation system. First, there is no shared understanding of what *policy evaluation* is, leading to a confusion from several actors about its goals and methods. Second, institutional mandates are not always clearly defined, meaning that the responsibility for who is to conduct evaluations is not clear. Finally, and perhaps most importantly, there is no high-level guidance orchestrating the policy evaluation system across government, which affects the quality and impact of evaluations.

In this context, this chapter discusses the extent to which the Portuguese government has institutionalised policy evaluation from a whole-of-government perspective and suggests some policy recommendations based on good practices identified in other OECD countries.

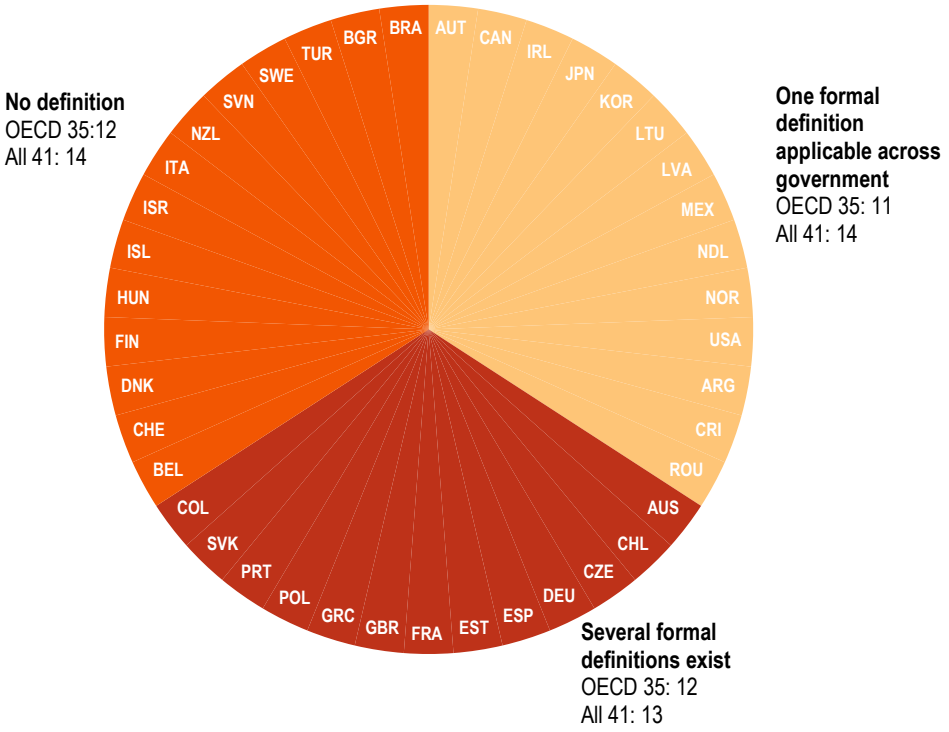
1.1. Creating a shared understanding of policy evaluation

In Portugal, there is no official definition of policy evaluation. Whilst evaluation is mentioned in some sectoral laws or regulations invoking the need to evaluate a programme, it is never defined explicitly (PlanAPP, 2022[3]). In practice, this has led to confusion from some government actors as to its goals and methods. Indeed, some actors use the term policy evaluation to refer to different practices in particular, assessments or monitoring.

Yet, empirical evidence suggests that a clear definition of policy evaluation represents an important starting point for governments to distinguish it from other practices and to create a shared understanding across the public sector of its aims, tools and features (OECD, 2020[2]). The OECD Recommendation on Public Policy Evaluation defines policy evaluation as "*a structured and evidence-based assessment of the design, implementation or impacts of a planned, ongoing or completed public intervention, its design, implementation or results* (OECD, 2022[1])." (see Box 1.2).

In 2020, 23 OECD countries had adopted at least one explicit definition of policy evaluation, while 11 countries have adopted a single shared definition across government (see Figure 1.1). Many of these countries embed their definitions can be embedded in official guidelines or in legal frameworks.

Figure 1.1. Definition of policy evaluation



Note: n=41 (35 OECD member countries).
Source: OECD (2020[2]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, https://doi.org/10.1787/89b1577d-en.

This is the case in Lithuania and the United Kingdom (see Box 1.2). Indeed, in the United Kingdom the definition of policy evaluation is contained in the main evaluation guidelines, the Green Book, while in Lithuania the definition is embedded in legal text.

Box 1.2. Definitions of policy evaluation in Lithuania and the United Kingdom

In the **United Kingdom** policy evaluation is defined as ""the systematic assessment of an intervention's design, implementation and outcomes. Both monitoring and evaluation should be considered before, during and after implementation" (HM Treasury, 2022[4]).

In **Lithuania** policy evaluation is defined as *"a systematic and objective determination of the suitability, effectiveness, efficiency, usefulness and long-term impact of the planned, executed or completed programs"*. Lithuania assigns to policy evaluation the quality attributes 'systematic' and 'objective' and states also the multiple criteria which must be evaluated. Once again, it does not offer specific guidance on the time setting and the actors involved (OECD, 2021[5]).

Source: (OECD, 2022[1]), (HM Treasury, 2022[4]) (OECD, 2021[5])

Establishing an official government definition of policy evaluation in Portugal could help clarify doubts around this practice and help differentiate it from other policy tools. As a relatively new agency aimed with the mandate to promote policy evaluations across government, PlanAPP could develop such a definition. In order to do so, it could work together with relevant stakeholders, in particular from the Agency for Development and Cohesion (*Agência para o Desenvolvimento e Coesão* in Portuguese, AD&C), as well as line ministries and any relevant external actors (academia for example). In order to actually promote a shared understanding of this practice, on the medium to long term, the definition could be included in an official public document, such as evaluation guidelines or a legal framework. As is the case in most OECD countries mentioned above, the Portuguese definition could include elements related to:

- quality attributions of this practice (e.g., the notion of a "structured and evidence-based assessment").
- when it should be conducted (e.g., *ex ante* or *ex post*).
- what types of public interventions are evaluated (policies, reforms, plans, laws, etc.).
- what is the object of the evaluation (design, implementation and/or results of the intervention).
- with what criteria (efficiency, effectiveness, impact, sustainability, etc.).
- the actors that are involved (e.g., internal or external evaluators).

1.2. Cross-government high-level guidance on policy evaluation

As policy evaluation requires significant resources and time, forward planning is important to organise evaluation practices across government, ensure that they are timely and fit-for purpose, as well as that policies are regularly evaluated. In this regard, high-level cross-government guidance clarifying when evaluators should conduct evaluations, how and for what purpose, is crucial. Such high-level guidance can take the form of evaluation agendas, as well as legal or policy frameworks.

In Portugal, there is no such cross-government guidance on when to perform evaluations or clarity on what policies need to be evaluated. In addition, there are currently no government-wide legal or policy frameworks in Portugal establishing rules on policy evaluation (PlanAPP, 2022[3]). As a result, evaluations are conducted on an *ad hoc* basis, practices are not homogenised across ministries, and there is little forward planning on evaluations, ultimately hindering their impact and use. One important exception are the evaluations of structural fund programmes, which are organised across government by the Agency for Development and Cohesion's (*Agência para o Desenvolvimento e Coesão* in Portuguese, AD&C) the evaluation agenda.

Across OECD countries, evaluation is embedded in different ways. A majority of countries have developed a legal framework that guides policy evaluation across government (OECD, 2020[2]). In addition to that, some countries have also developed policy framework (15 OECD countries) (OECD, 2020[2]). Finally, some countries have adopted evaluation agendas to create a cross-government plan for conducting and using policy evaluations, as is the case in the Netherlands (see Box 1.3). Evaluation agendas can provide high-level guidance to help line ministries think early about their evaluation needs and potential uses for these evaluations. They also promote transparency in line ministries' decisions to supply evidence.

Box 1.3. The Netherlands' Strategic Evaluation Agenda

The Netherlands adopts a multi-annual **Strategic Evaluation Agenda (SEA)** in order to better systematise evaluations and ensure that their impact decision making is increased. This new instrument has proved an interesting way to better plan and keep track of evaluation practices across the ministries, as well as to link policy evaluations and evidence to budgetary processes.

The SEA is developed on the basis of discussions between the line ministries and the Ministry of Finance and involves regular meetings between key ministerial stakeholders over the budget year. This is to ensure both accountability and policy learning. In addition, the SEA is the same duration as the government mandate and it is shared with Parliament to ensure that it aligns to the government agenda.

Each ministry also publishes their own SEA as part of their yearly budget memorandum and covers the entire policy/evaluation cycle: the *ex ante*, the *ex durante* and *ex post* evaluations. The evaluations that are planned in this way are then divided by policy theme and are presented together with the current knowledge base and knowledge gaps for each theme.

The Strategic Evaluation Agenda was created to overcome previous bottlenecks that the Dutch evaluation system was facing. In particular, every policy had to be evaluated 7 years after it was adopted. This system was found to be excessively rigid, leading to evaluations being conducted as "check-the-box" exercises, which ultimately hindered the impact of the evaluations. Indeed, performing an evaluation after 7 years meant that the evidence came at a time that did not necessarily match decision makers' evidence needs well, thus hindering the possibility for evaluations to generate policy improvements. As a result, the SEA was put in place to allow line ministries and the Ministry of Finance to plan evaluations based on when they would be most needed for accountability, learning and policy development.

Source: (Ministry of Finance, 2022[6])

Adopting such an evaluation agenda could be a first step in better systematising policy evaluation in Portugal. In this context, PlanAPP could consider developing a multi-year evaluation agenda, in consultation with line ministries, to establish high-level guidance on policy evaluation over a medium-term period. This planning document could clarify the policy evaluations that need to be conducted across the executive over the next few years. The purpose of developing an evaluation agenda is to promote deliberate planning of evaluation activities in order to ensure that results are available when they are most likely to be used and to improve data collection. It is also an important communication instrument towards decision makers and stakeholders, and thus a way to maximise the impact of evaluation results.

In developing such an agenda, PlanAPP could help identify the evidence needs of line ministries and be in charge of developing the agenda in collaboration with them. This process should first rely on the identification of any potential evidence gaps. In addition to the evaluations that line ministries plan to conduct, PlanAPP may wish to identify strategic policy areas that require co-operation across ministries where it could play a role in conducting its own evaluations (or co-ordinating the work of several entities). PlanAPP may also wish to consult some key external stakeholders in the process. Such an evaluation agenda would benefit from covering a medium-term period (e.g., 3 or 5 years) and be updated annually to remain current to policy priorities. To do this, PlanAPP could schedule an annual conversation with evaluation departments or units in line ministries, to get feedback from the ministries on upcoming their evidence needs.

On the longer term, in addition, Portugal should consider adopting a legal or policy frameworks to ensure that evaluations are conducted systematically regardless of changes in governments. A legal framework would be beneficial in order to underline the most basic and essential characteristics of the evaluation system: a definition of policy evaluation, the frequency and timing of evaluations, and the different mandates and responsibilities inside the system (see next section for more information on this latter point). Even if legal frameworks are not, in of themselves, a guarantee of success, they can considerably help the systematisation of policy evaluation – particularly in a country with a civil law culture such as Portugal.

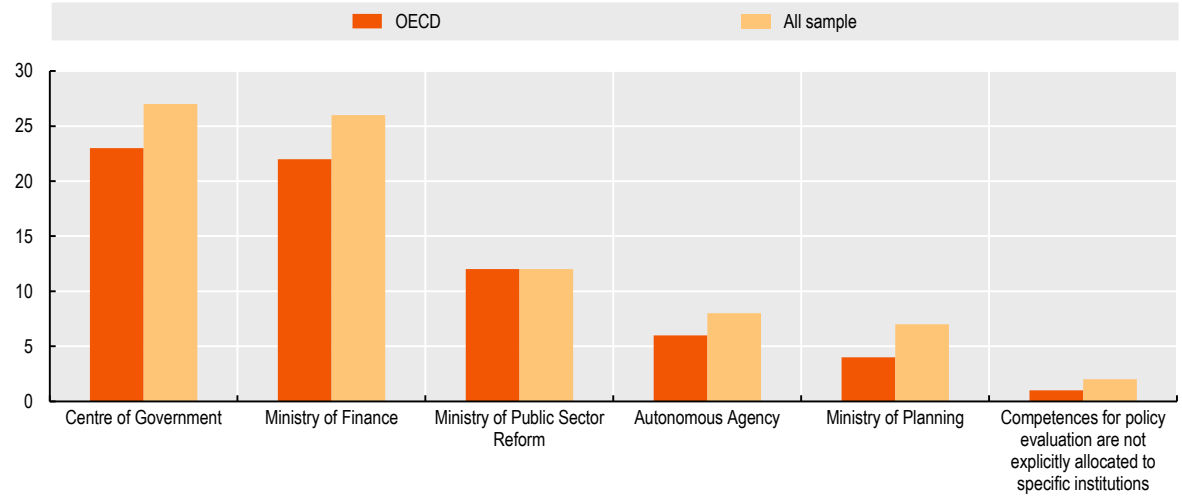
1.3. Attributing institutional mandates inside the executive

In order to institutionalise policy evaluation, it is important to clarify who does what. For this reason, legal frameworks or policy frameworks generally establish what institutions carry out policy evaluation and how they co-ordinate with each other. This is particularly useful in contexts where there are multiple institutions with different functions and roles, in order to clarify what is expected from each organisation.

In Portugal, not all institutional mandates are clearly defined. First, the recent decree n.21/2021 on "The approval of the organisational structure of the Competence Centre for Planning, Policy and Foresight of the Public Administration" gave Plan APP the mandate to co-ordination policy evaluations across government (Presidency of the Council of Ministers, 2021[7]). In addition, AD&C does have a clear mandate in co-ordinating the evaluation of programmes funded by the European Union structural funds, as established by the law decree n.140/201 (Presidency of the Council of Ministers, 2013[8]). On the other hand, few ministries clearly have policy evaluation as part of their mandate (that is, that policy evaluation is mentioned explicitly in their organisational decree for instance). As a result, policy evaluations are conducted on an *ad hoc* basis.

Yet, having such clear mandates is important to ensure that institutions clearly allocate time and resource to policy evaluation and to make sure that actors know what their role is in the evaluation system. Across the OECD, institutions within the executive often times have two different types of mandates: either the mandate to conduct policy evaluation (as is the case for line ministries, for example) or that to co-ordinate evaluations across government and to share best practices. While there is significant variation across OECD countries as to what institution takes on what role, evidence suggests that co-ordinating evaluations across government is most often the responsibility the Ministry of Finance, the Prime Minister's Office, or an autonomous agency (OECD, 2020[2]). The choice to allocate this role to an institution or another may depend on the institutional landscape of a country, as well as the main reason behind conducting evaluations (OECD, 2020[2]). For instance, programme evaluation for budget decision making may be led by the Ministry of Finance, whilst evaluations of structural funds is often led by an autonomous agency.

Figure 1.2. Institutions within the executive that have competences related to policy evaluation across government



Note: n=42 (35 OECD member countries). Answers reflect responses to the question, "Which of the following institutions within the executive have competences related to policy evaluation across government? (Check all that apply)". Answer option "other" is not displayed.
Source: OECD (2020[2]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences,* OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/89b1577d-en.

In the Netherlands and in the United Kingdom, both the mandates to co-ordinate evaluations across government and that to conduct (or commission) evaluations are clearly attributed. Indeed, in the Netherlands, line ministries are the main actors responsible for conducting *ex ante* or *ex post* evaluations. Each line ministry has a Directorate for Financial and Economic Matters (FEZ), which, among other tasks, supervises the evaluation on effectiveness and efficiency of policies. The Bureau for Strategic Analysis (BSA) inside the Ministry of Finance plays a general co-ordination role for evaluations and supervises line departments. It also manages the Strategic Evaluation Agenda. As such, the Ministry of Finance is responsible for overseeing the evaluation policy of all departments and provides assistance to line ministries for their evaluation work. Similarly, in the United Kingdom each line department is responsible for conducting evaluations. In addition, the Evaluation Task Force (ETF), which sits between the HM's Treasury and the Cabinet Office, also has the mandate to promote good practices across departments and to ensure the quality of evaluations.

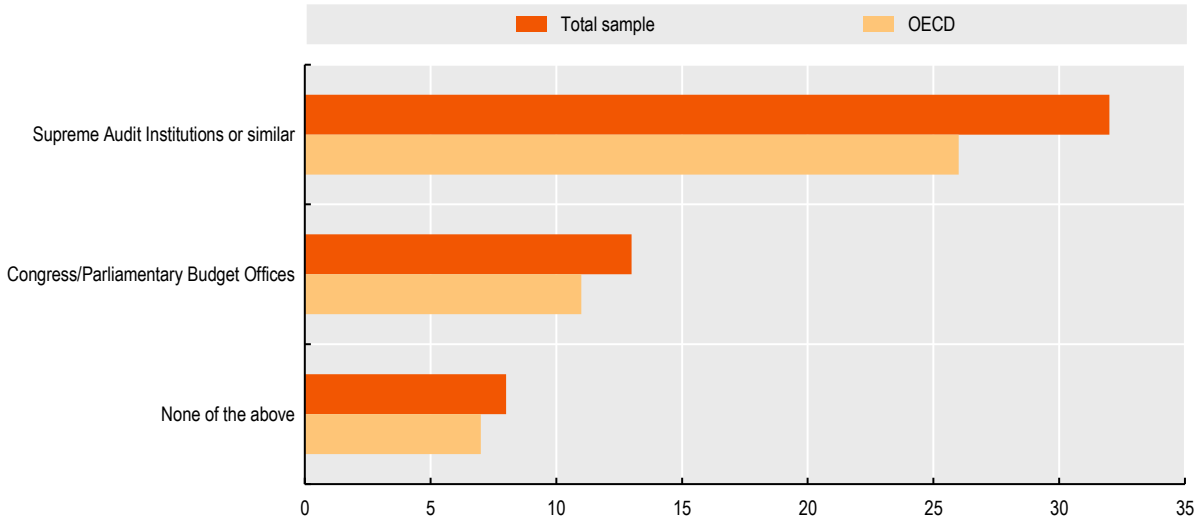
In Portugal, line ministries' mandates to conduct policy evaluations could be further clarified. A first step in clarifying the roles in the evaluation system would be to assign line ministries an explicit mandate to conduct policy evaluations. In the absence of a cross-government legal framework, line ministries could consider including a mandate to conduct policy evaluations in their organisational decrees. This will be important in order for the Portuguese government to ensure that all sectorial policy topics are evaluated. Whilst giving each line ministry an evaluation mandate is a first important step, another important aspect will be to clarify, internally to the ministries, which units are responsible and what resources are dedicated to this activity. Beyond the legal aspect of attributing such a mandate, on the longer term, it would also be useful for line ministries to clarify which teams or units are responsible for conducting evaluations, as well as what resources (human and financial) are dedicated to this task. This will be particularly important in order to determine who, inside the line ministry, will be PlanAPP's main contact point on issues that fall under this new agency's mandate.

1.4. Promoting a rich evaluation system outside the executive

The OECD Recommendation on Public Policy Evaluation encourages governments to *"foster a culture of learning and accountability by promoting demand for, and ownership of, evaluations within and beyond the executive"*. Indeed, having a number of actors performing or using evaluations outside of the executive helps promote an overall culture of evaluation, that is, a culture in which evaluation results achieve impact.

Across OECD countries, there are a number of institutional actors performing evaluations outside of the executive (OECD, 2020[2]). Evidence suggests that the most important role is played by Supreme Audit Institutions (SAIs) and Parliaments, in particular by the Parliamentary Budgetary Offices (PBO) (see Figure 1.3). Supreme Audit Institutions and Parliaments either conduct, commission or discuss evaluation results. Indeed, these are often important evaluation "consumers" which can promote the use of evaluations in the system and, ultimately, evidence-informed decision making. In some cases, SAIs even play a role in promoting the quality of evaluations by issuing guidelines. A good example in this regard comes from the UK where the NAO and the Parliament play an important role in policy evaluation (see Box 1.4). A more modest but emerging role is also played by the French Parliament to increase the use of policy evaluations (see Box 1.4).

Figure 1.3. Actors outside the executive with role in policy evaluation



Note: n=42 (35 OECD member countries). Answers reflect responses to the question, "Which of the following institutions beyond the executive have competences on policy evaluation at the central/federal level? (Check all that apply)".
Source: OECD (2020[2]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, https://doi.org/10.1787/89b1577d-en.

Box 1.4. The role of actors outside the executive: France and the United Kingdom

In the **United Kingdom**, the Parliament has an important role in the field of evaluation and it often demands evaluations through its various parliamentary committees (Mathot and Giannini, 2022[9]). Another essential actor is the National Audit Office (NAO) which performs Value for Money reports and often assess the quality of the evaluation system in the government. The Value for Money reports are presented and used in Parliament and they often involve the call of government officials to answer questions regarding the implementation of policies and programmes and the efficiency of public spending.

In **France**, in the recent years the Parliament has gained significantly more interest in policy evaluation. In this context, France Stratégie, together with some departments in line ministries, has organised the *Printemps de l'évaluation* or Evaluation Spring in Parliament. The Evaluation Spring in France was a two-day event organised once a year where different line ministries would present the results of evaluations to members of Parliament. In 2021, 46 evaluations were presented to members of parliament.

Source: (Mathot and Giannini, 2022[9])

Other actors at arm's length of the executive, such as knowledge brokers (which can be think tanks, research institutes, semi-public agencies, etc.) promote the impact of evaluations by synthesising and communicating their results.

In Portugal, there are still relatively few actors outside of the executive that are really engaged in conducting policy evaluations, discussing them or communicating their results. One important exception is the Court of Accounts, which has recently invested in improving its analytical capacities, but does not yet perform policy evaluations.

In order to promote the external role of the Court of Accounts and of the Parliament as users of evaluation results, PlanAPP could consider sharing its multi-annual evaluation agenda with these actors. Sharing the agenda with the Court of Accounts would be a way for the government to share information with the Court on where it is concentrating its evaluation efforts. It could also be an opportunity to avoid duplications between the work being done by the Supreme Audit Institution and PlanAPP, without prejudice to the independence of the Court.

Second, sharing PlanAPP's multi-annual evaluation agenda with the Parliament could be a way to foster interest in evaluation results. In addition to ensuring greater impact of PlanAPP's evaluations, this could ultimately create an incentive for line ministries to perform policy evaluations, in so far as they know that they will more likely be used, or at least, read. To further foster the Parliament's role inside the Portuguese evaluation system, PlanAPP could organise some yearly dissemination events in collaboration with the Technical Unit for Budget Support (*Unidade Técnica de Apoio Orçamental* (UTAO)) on evaluation, as well as trainings for members of Parliament.

2 Promoting quality of policy evaluations

Quality of evaluations is essential to ensuring their use for decision making, and thus for evaluations to serve as tools for learning, accountability and better decision making. However, quality is an important challenge faced by policymakers and practitioners in this area. This is due to a mix of capacity and skill gaps, issues with access and usability of data, heterogeneous oversights in the evaluation processes, and insufficient mechanisms for quality control (OECD, 2020[2]). Yet, insofar as good quality evaluations benefit from greater credibility, they are likely to be given more weight in decision making. The issue of the quality of evaluations is therefore a crucial one.

The OECD Recommendation on Public Policy Evaluation underlines that countries need four key elements to promote the quality of evaluations (OECD, 2022[1]):

- **Engage with stakeholders** and intended users to ensure that evaluations can be used.
- Build **public sector skills for evaluation** by conducting regular training, recruiting and retaining employees with the adequate skills or collaborating with academia, the private sector and other jurisdictions to improve the availability of these skills.
- Ensure the **availability of high quality, timely, accessible, disaggregated and re-usable results,** performance and administrative data for policy evaluation.
- Establish **quality standards and mechanisms for evaluations** to generate robust and credible evaluation results that can be trusted and used with confidence.

In Portugal, capacities and skills for policy analysis seem to be lacking in the public sector. The supply of skills for evaluation is overall low in the Portuguese marketplace. As of 2022, there were only to university programmes focusing on policy analysis (PlanAPP, 2022[3]). In addition, access to high quality and re-usable data is challenging for line departments in ministries. Line ministries are often unable to access data that is useful to perform evaluations. Finally, there are no shared and formalised quality assurance mechanisms in place to ensure that evaluations meet high quality standards in terms of governance and of methodological robustness. All these elements significantly hinder the final quality of the evaluations and the trustworthiness of the results.

This chapter discusses the extent to which the Portuguese evaluation system is able to ensure the quality of evaluations. The chapter also identifies good practices from other OECD member countries that could be of inspiration to Portugal. Finally, it provides concrete recommendations to Portugal on how to promote the quality of its policy evaluations.

2.1. Engaging with evaluation stakeholders and users

Policy evaluations often benefit from the viewpoint of different actors. These are often academics and experts but also administrators, businesses and citizens that are impacted by the policy. For this reason, involving different stakeholders at crucial phases of an evaluation can enhance transparency in the process

and improve data collection. Such engagement can also increase the impact of evaluations by allowing stakeholders to give feedback on evaluation questions and recommendations, thus improving their usefulness.

Systematic stakeholder engagement does not take place for all evaluations in Portugal and practices vary significantly across line ministries and agencies. The AD&C agency does involve several actors in the evaluation of its programmes and does so from early stages of the evaluation process in the so called *Grupo de Acompanhamento*. On the other hand, many line ministries do not involve any stakeholders in the evaluation process.

The OECD Recommendation on Public Policy Evaluation recommends countries to involve stakeholders from early stages of the evaluation cycle. An interesting practice in this regard is that of France Stratégie in France (see Box 2.1).

Box 2.1. The governance of France Stratégie's policy evaluations

France Stratégie involves several actors in the evaluation process making sure that knowledge coming from different sources, disciplines and expertise is mobilised and well-integrated in the final product. For each policy area in which France Stratégie wants to perform an evaluation, a special Commission is involved. This commission is composed of stakeholders and experts that can support the internal staff of France Statégie in performing the evaluations. There are currently 11 standing Commissions at France Stratégie. These are generally composed of social partners, representatives of the public administration, Parliamentarians and academics. The commissions are chaired by France Stratégie. They intervene at different moments in the evaluation process, such as by helping ensure proper planning of data collection efforts.

When it comes to the policy evaluations that PlanAPP conducts itself, the agency should strive to systematically engage with relevant stakeholders throughout the process, in order to create ownership for change and trust in the evaluation results. In particular, PlanAPP could establish a standing network of actors to engage with. This group of stakeholders could function as the advisory group that was suggested earlier, and bring together people from universities, as well as research centres and civil servants. Together with these institutional stakeholders, some public consultations could also provide an interesting approach to promoting the impact of evaluation results.

2.2. Fostering capacities and skills for policy evaluation

Capacities for policy evaluation refer to the resources, both in terms of people and in terms of budget, dedicated to the activity of evaluations. Human resources are, together with financial resources, key enablers for quality policy evaluations (OECD, 2020[2]). Indeed, conducting policy evaluations can be financially and labour intensive. As a result, they require a critical mass of capacities inside public administrations in order to design evaluations, collect and analyse data, as well as disseminate their findings. Together with enough resources, it is also important for public administrations to have the right skills to conduct policy evaluations. Indeed, evaluation requires specific skills. Both capacities and skills play an essential role in ensuring the quality of evaluations and, ultimately, increasing their use.

In Portugal, responsibilities for conducting policy evaluations are often spread out across different units and departments in a given ministry, such that it can sometimes be difficult to identify exactly who is responsible for conducting evaluations. In addition, where evaluations are conducted by one single unit in a ministry, the team often has several tasks assigned to it, such as conducting statistical analysis, planning and foresight, or international relations. This might make it difficult for staff members to dedicate time to conducting

evaluations, particularly given that policy analysis is a longer-term task that may be disrupted by the more short-term (or even urgent) demands related to some of the other responsibilities previously mentioned.

Moreover, human resources and skills dedicated to policy evaluation are scarce inside the Portuguese public administration. Indeed, the public administration is unable to attract and retain these skills for several reasons. First, evaluation skills are quite rare in the Portuguese job-market. In addition, the Portuguese civil service exams do not earmark profiles to specific jobs, which means that administrations with specific skill requirements are not able to find the right profiles for their needs. Finally, wage and career opportunities are not attractive enough in the public administration to attract these types of competences which are also in high demand in the private sector.

As the quality of an evaluation is closely linked to that of the evaluators, it is essential for governments to find a way to attract and retain these skills. This is one of the main challenges for public administrations in all OECD countries (OECD, 2020[2]). Indeed, when asked to rank the challenges for the promotion of policy evaluation, the lack of capacities and capabilities of human resources was ranked third (OECD, 2020[2]).

However, countries are adopting different strategies to ensure that the public administration has sufficient capacities and skills for evaluation. Some countries have developed specific career paths to ensure a sufficient number of civil servants with analytical skills. This is, for example, what is done in the United Kingdom (see Box 2.2). Another important way to make sure these skills are available is by building them through in-house through more structured training programmes for public servants. This is the case of France, for example, where the National School for Statistics and Economic Administration has developed master programmes to train future economists and statisticians in policy analysis (OECD, 2021[5]). Other countries are also developing one-off trainings to foster internal skills. An example is Lithuania which has recently developed a training on policy evaluation in STRATA. Indeed, the Government Strategic Analysis Centre (STRATA), the Lithuanian evaluation champion, has developed some in-house trainings for evaluators. These are structured as four half-days (from 9:45 to 15:00) courses covering several aspects from the definition of policy evaluation to the use of evaluation results (STRATA, 2022[10]). In 2022, STRATA ran this training five times reaching more than 100 civil servants.

Box 2.2. The United Kingdom approach to analytical skills inside government

Inside the UK public administration there is a considerable number of professionals with analytical skills (around 15 000 policy professionals) (OECD, 2021[5]). There are different specific career tracks inside the public administration for people performing evaluations. Moreover, these professions have access to a support unit called the Government Economic and Social Research Team, in HM's Treasury. This team offers support across government departments, it co-ordinates learning and development activities, maintains and develops profession standards, and provides both formal and informal mechanisms for developing standards for evaluation. Overall, the UK government supports these professions in a series of ways:

- **Trainings**: both on specific methodological aspects and on cross-cutting themes. For example, the Business, Energy and Industrial Strategy department (BEIS) offers all employees in the evaluation unit a three-day training that is organised by the UCL (University College London).
- **Informal networks:** these can be across government (e.g., the Cross-Government Evaluation Group), connecting people from all the government to share their experience and knowledge, but also internal to specific departments. For example, the BEIS has an Evaluation Practitioners Network (EPN) that is composed of internal employees of the BEIS working on evaluations. Networks are particularly useful to foster the skills for use of evidence, as these skills require less technical expertise with respect to the analytical ones which might require also longer trainings.

- **Advisory Panel for Evaluation:** The Cross-Government Trial Advice Panel is a free-to-use service that provides government departments with support for designing and implementing effective trials and evaluations.

Source: (OECD, 2021[5])

To increase the availability of analytical skills, the Portuguese government should consider investing in trainings for civil servants in public policy analysis and evaluation, as well as consider creating a masters-level programme in this field of study. In other countries, public administrations develop trainings in co-ordination with national universities to upskill civil servants. In Portugal, PlanAPP could similarly support the Institute of Public Administration, Instituto Nacional de Administração (INA) in the development of such a training course.

Other than specific policy analysis and evaluation skills, evaluation champions also require a multi-disciplinary team to carry-out the different tasks assigned to them. Indeed, evaluation champions can have vastly different responsibilities which require different skill sets: they may facilitate the use of evaluations from decision makers through communication efforts, help line ministries with the quality of their evaluations through quality assurance and control, or even conduct cross-sectorial evaluations. In the case of PlanAPP, which aims to position itself as a policy evaluation champion, the Evaluation Unit (Unidade Técnica de Avaliação (UTA)) would need to bring together some communication/knowledge dissemination skills, as well as more advanced evaluation expertise. The formers are essential to create tailored products (like policy briefs, infographics, user-friendly databases etc.) that capture the attention of policymakers and make evaluation results easy to understand. Evaluation experts are important both to conduct internal evaluations, as well as to understand and support line ministries with their evaluation activities. Together with these more technical skills, PlanAPP could benefit from networking and engagement competences, to support their role in co-ordinating a network of government evaluators and communicating with stakeholders inside and outside the public administration. In order to make sure that it has access to these different types of skills, PlanAPP should consider investing in the Evaluation Unit's skills in the medium-term by hiring staff members with more advanced skills in evaluation and evidence synthesis. The Evaluation Unit may wish to rely on PlanAPP's communication unit for help with its responsibilities related to the dissemination of evaluation results, or better yet, hire a small number of staff members with these skills and a lay understanding of evaluation methods.

Finally, another way to foster skills for evaluation inside the Portuguese public sector is to tap into the expertise and knowledge of academia. First, PlanAPP and the Portuguese government overall could consider hiring academics on secondment to work on specific policy evaluations, and more generally to access technical skills that are not present in the public sector. On the longer term, PlanAPP would greatly benefit from developing a network of evaluators, including academia, to promote the quality of government evaluations and create a space for exchange on related topics. For PlanAPP specifically, creating a network can provide support to the development of methodological guidelines and trainings for line ministries. The academics present inside the network can also work as peer reviewers of evaluations ensuring that the end-product respect some quality requirements. For the Portuguese government, in general, such networks are beneficial as a way to increase the evidence/science to policy interface, and thus promote the supply of relevant evidence for policymaking and, conversely, increase the impact of scientific and research evidence.

2.3. Availability of, and access to, quality data

High quality and timely data are essential to producing reliable and robust analyses. In order for data to meet the quality criteria to be used for evaluation, it needs to be accurate, verifiable and documented

(OECD, 2020[2]). In addition, accessing different sources of data is essential to ensuring that departments can unlock insights for policy evaluation. These sources can be big data, open data, statistical data and they all contribute to enriching the results of the evaluation. The lack of high-quality data can significantly hinder policy evaluation.

In Portugal, line ministries perceive data access as one of the most significant barriers to policy evaluations as several analytical units are unable to access high quality data for evaluation purposes. This is because the statistical law of 2008 establishes that access to micro data should be granted only to researchers for scientific purposes (Article 6 paragraph 7-8). The protocol of May 2022 signed by INE and the Directorate-General for Education and Science Statistics, (in Portuguese, the Direção-Geral de Estatísticas da Educação e Ciência, DGEEC) specifies what entities are considered "researchers" and what purposes can be considered "scientific". This offers a clearer picture of who can access microdata (INE, 2022[11]). According to this protocol, one of the available options to access microdata is to be a research entity recognised by Eurostat on the date of data access (INE, 2022[11]). In Portugal, at the moment, almost all these recognised entities are universities. The only exception inside the government is represented by the Office of studies and international relations (Gabinete de Planeamento, Estratégia, Avaliação e Relações Internacionais, in Portuguese, GPEARI) (Eurostat, 2022[12]). For this reason, line ministries still encounter significant problems in accessing microdata and must rely on the presence of an accredited researcher inside the unit to perform evaluations internally. However, these skills are rarely present inside the public administration as it was previously mentioned.

Across OECD countries, data related issues are considered important barriers to policy evaluation. Overall, the main data challenges can be summarised as:

- a lack of available data;
- issues with data access;
- capacity gaps among government departments and agencies to generate data in a format that can be used.

To tackle these issues countries have adopted several approaches to ensure the secure access to the relevant data when performing evaluations. A first important way to address the data issue is to think early about data needs. In the United States for example data needs are expressed in advance through the learning agendas, which are developed by each department to reflect on the evidence needed for the following year.

Data access for evaluation purposes is different across countries. In some countries, like France most line ministries are equipped with a statistical unit and face lower barriers to accessing data in the public administration (see Box 2.3).

Box 2.3. Public entities that perform research activities in France

The central public administration is the one of the actors producing policy evaluations in France (Baïz, 2022[13]). To ensure this function it was essential to equip evaluation units with the high-quality data needed to perform such evaluations. In this respect, with the Law for a Digital Republic from 16 October 2016 (*Loi pour une République numérique*) access to microdata was granted to all public agencies performing research activities (article 36). This condition (performing research activities) is easier for public administrations to prove than demonstrating that they are a research entity recognised by Eurostat.

Source: (Assemblée nationale, 2016[14])

To grant access maintaining high standards of security, other countries have instituted secure data access points from which evaluators can access microdata for evaluations. This is for example what the Netherlands has done (Box 2.4).

Box 2.4. Ensuring data access in the Netherlands

The **Central Bureau of Statistics** in the Netherlands provides access to data following request from the line ministries. This is regulated by the Statistics Netherlands Act of 2003 which in articles 41 allows the CBS to provide or grant access to a set of data to "*research departments of ministries and other departments, organisations and institutions, in so far as the Central Commission for Statistics (CCS) has given its consent*" (Dutch Parliament, 2003[15]). Hence, CBS often interlinks databases and provides microdata to the ministries. Moreover, the CBS is equipped with a Remote Safe Centre which allows it to share data safely through Remote Access (RA). A research commissioned by the CBS and performed by independent researchers has highlighted that the Remote Access granted by the CBS complies with all the security and privacy standards (Berg et al., 2020[16]). For this reason, line ministries in the Netherlands usually do not experience particular problems in accessing data for evaluations.

Source: (Dutch Parliament, 2003[15]) (Berg et al., 2020[16])

PlanAPP should also work to facilitate data access for line ministries. An initial option would be that line ministries seek to get a research entity status as is the case for PlanAPP. However, a smoother path should be available to line ministries on the longer term. In France, for example, if France Stratégie is recognised by Eurostat as a research entity, not all evaluation units in line ministries benefit from the same status. Yet, they do not encounter particular issues in accessing microdata, as France allows public administrations to access microdata for research activities. On the longer term, the Portuguese government should therefore consider amending the statistical law of 2008 to make it explicit that evaluation and analysis units inside line ministries are one of the categories of actors that can access micro data.

Portugal, as all OECD countries, is confronted with the challenge of balancing the use of personal data for evidence-informed decision making and ensuring that the personal data rights of citizens are secured and respected. Finding an appropriate solution is essential as both policy learning and citizens privacy must be guaranteed (OECD, 2021[5]). In this context, several countries have developed data governance strategies to ensure that data can be accessed for research purposes maintaining high levels of privacy. This is the case of the United States which in 2018 has adopted the Foundations for Evidence-Based Policymaking Act. The Evidence Act aims for federal agencies to better acquire, access, and use evidence to inform decision making. (OECD, 2020[2]). An important role in the implementation of this data strategy is played by the annual evaluation plans which are also useful to strategically plan which data needs will emerge across the executive.

The development of an analogous document in the Portuguese public administration could be beneficial to shed light on the role of line ministries in providing data for evidence-informed decision making and in how this role can be granted. The state secretary for Digitalisation and Administrative Modernisation could work on such a plan together with PlanAPP.

2.4. Ensuring good governance and technical quality of evaluations

To be robust and trustworthy, evaluations need to be independent and methodologically sound. For this reason, governments can put in place mechanisms to ensure good governance and methodological robustness of policy evaluations. These can be either quality assurance mechanisms, which seek to ensure credibility in how the evaluation is conducted (the process), either quality control mechanisms which instead focus on the end product (the report) ensuring that it meets a certain standard for quality (OECD, 2020[2]) A mix of quality assurance and quality control mechanisms is very important to really foster the quality of evaluations. They are also important to ensure some comparability of evaluations across ministries.

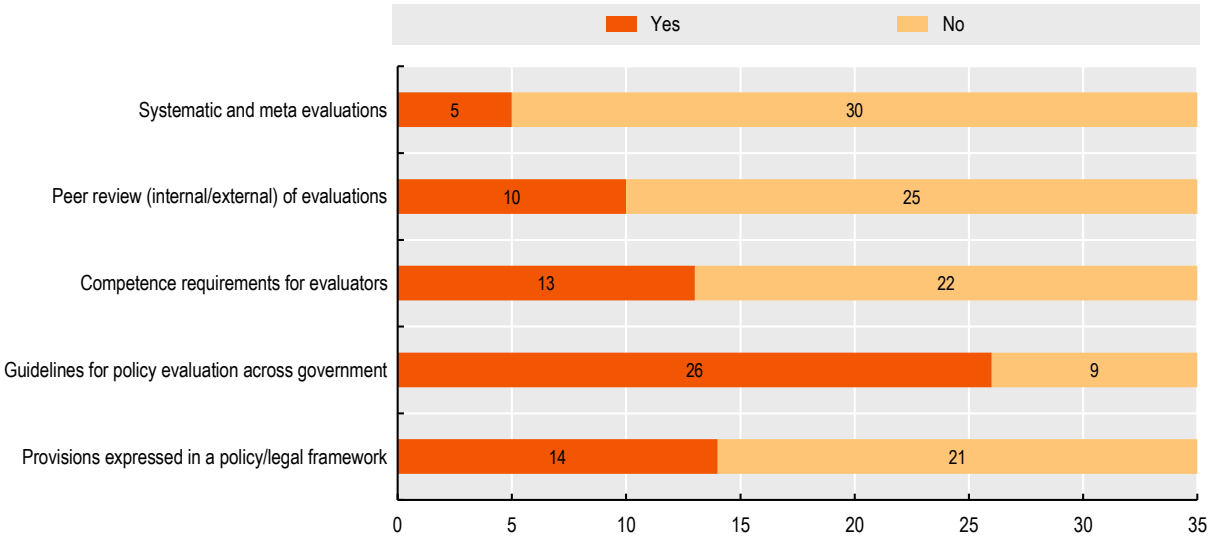
These mechanisms are still not in place in the Portuguese system and evaluations are conducted in different ways by each line ministries. For this reason, evaluations vary a lot in terms of structure, content and quality across government and monitoring reports are sometimes confused with evaluations.

The OECD Recommendation on Policy Evaluation provides significant direction on how to ensure the good governance and technical quality of evaluation process:

- developing standards on the quality of the evaluation process, which can be embedded in evaluation guidelines or in legal/policy frameworks.
- controlling the quality of the evaluation product.
- fostering quality at an institutional level.

Given the different instruments available to promote quality from a technical and governance perspective, countries have developed very different quality assurance and control systems (Figure 2.1). These usually imply the use of multiple instruments in order to cover both the technical and the governance aspects. Among these instruments, guidelines are the most commonly used (see Box 2.5 for some examples). These can address both technical quality and good governance (17 OECD countries out of 35 have them). Other instruments to ensure the technical rigour are networks of experts and peer review processes.

Figure 2.1. Mechanisms used for the promotion of quality in OECD countries



Note: This chart refers only to the 35 OECD countries that responded to the Policy Evaluation Survey (2019)
Source: OECD (2020[2]), *Improving Governance with Policy Evaluation: Lessons from Country Experiences,* OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/89b1577d-en.

Box 2.5. Different guidelines on policy evaluation

In **France**, there are different methodological guidelines to foster the quality of evaluations, which were created by France Stratégie in 2016 and 2017 and are respectively on *How to evaluate the impact of public policies?* and *Guideline to socioeconomic evaluation of public investments in France*. The latter offers specific instructions on public investments which have some peculiarities with respect to other public policies. Together with these guidelines, single departments have developed their own policy-specific evaluations guidelines. More recently, to also provide insights on different methodologies, France Stratégie has developed, together with the Interdisciplinary Laboratory for the Evaluation of Public Policies (LIEPP) an additional guideline looking both at qualitative and quantitative methods for policy evaluation (France Stratégie, 2022[17]).

In **Lithuania,** guidelines are the main instrument to ensure quality of policy evaluations across government. One of the most complete documents in this regard is the *Guidance on the application of the methodology for evaluating programmes* (2011) which gives advice on how to plan and design an evaluation, from identifying the need for an evaluation to establishing an evaluation plan, including methods (Ministry of Finance, 2011[18]). The Government resolution No. 308 adopted the Methodology for *ex post* evaluation of existing legislation providing information on how to prepare, how to carry out and how to publish evaluations (Government of the Republic of Lithuania, 2021[19]). It also highlights other more technical and methodological documents that should be used like the Methodological Guidelines for Social Cost-Benefit Analysis.

The Ministry of Finance, in **the Netherlands**, provides a Guidance on Policy Reviews to line ministries. These guidelines contain information both on the governance and the methodologies to be used when performing evaluations. Moreover, the Ministry has developed the Policy Evaluations toolbox which divides the process of carrying out a policy evaluation into five steps (Ministry of Finance, 2022[6]). Each step refers to concrete 'tools' (evaluation and data collection methods) and examples of evaluations. The second mechanism to ensure quality is through the use of at least one independent expert. This aspect has been included in the recent amendments of the Regulation on Periodic Evaluation.

In **the United Kingdom,** guidelines provide information on how the evaluation should be structured, the methods that should be used, the questions it should answer. The main guidelines are the **Green Book** and the **Magenta Book**, providing information also on quality standards as well as some ethical considerations that the evaluators must follow. Together with these main guidelines, several departments have also adopted additional guidelines to better support the use of different methodologies or approaches that can be different.

Source: (France Stratégie, 2022[17]) (HM Treasury, 2022[4]) (Ministry of Finance, 2011[18]) (Government of the Republic of Lithuania, 2021[19])

Together with guidelines, there are other additional mechanisms to ensure the quality of evaluations. Some mechanisms focus on the evaluators themselves, like adopting professional or ethical standards, while other mechanisms are more interested in the process in which evaluations are conducted. Such mechanisms want to ensure that evaluations are conducted independently or that they are reviewed by external experts for example through peer review systems. These can be structured process or more informal processes taking place through networks. Indeed, networks as well represent important mechanisms to build capacity and ensure quality (see Box 2.6).

Box 2.6. The role of networks to strengthen evaluation quality in the United Kingdom

Together with the guidelines, networks can also be considered a form of quality mechanism. They are instruments that allow for exchange and capacity-building. They can often also work as support mechanisms to resolve methodological issues and overall increase the quality of the final evaluations. As already mentioned, the UK has a cross-governmental network called the **Cross-Government Evaluation Group (CGEG)**. The network is composed of representatives from different evaluation units in government departments. Finally, peer reviews are essential to ensure transparency, rigour and independence of the findings. Each evaluation must be submitted to generally two reviewers and their comments must be published together with the final evaluation in order to make it possible for external readers to assess the extent to which the final evaluation reflected the reviewers' suggestions.

The development of different tools to ensure the good governance and technical quality of evaluation is essential in Portugal. Guidelines represent the most common quality assurance mechanism and are essential to ensure that evaluations are conducted using pre-defined quality standards. They are also a way to promote a common methodology and facilitate standardisation of the practice. PlanAPP should elaborate on the medium term, guidelines for policy evaluations to ensure the quality of line ministries' evaluations. In doing so, it can be inspired by the guidelines developed by other evaluation champions such as the HM treasury in the United Kington, STRATA in Lithuania or France Stratégie in France. To make sure that they are used, relevant and accepted by stakeholders, the guidelines would benefit from being developed in a collaborative manner with line ministries and external stakeholders.

Another way in which PlanAPP can promote the quality of the process and end product of evaluations is, as mentioned previously, by establishing a network with academics to foster capacity inside the public administration and to create some quality control mechanisms. Indeed, control mechanisms are important to ensure that the evaluations are conducted properly and that meet the pre-determined standards.

The most common control mechanism is the peer review process which consists of a panel or reference group, composed of external or internal experts, in charge of reviewing the technical quality of an evaluation (OECD, 2020[2]). The peer review process helps determine whether the evaluation meets the adequate quality standards and can therefore be published.

In this area, PlanAPP should establish an advisory group/ peer review group to control and provide feedback on policy evaluations conducted by line ministries. The advisory group should be composed of academics specialised in the field of evaluation as well as in the policy area that is being evaluated. In addition, as will be mentioned in the following section, the advisory group could include some key stakeholders from outside government where relevant (social partners for example). The advisory group's composition could therefore be adapted depending on the policy or intervention that is being analysed.

3 Conducting policy evaluations that impact decision making

Once evaluations are produced, they need to be used by policymakers, civil servants, practitioners, etc. Ensuring that policy evaluations impact decision making represents the last but essential step. Indeed, one of the most fundamental rationale for conducting policy evaluations is their usefulness in informing policy and decision making, in general, and improving the intervention they consider, specifically. The principal goal of evaluation is to support decision making with useful insights on public issues and evidence on the impact of policies and their underlying change mechanisms. For this reason, it is essential to facilitate the use of evaluations inside the government.

In order to conduct public policy evaluations that impact decision making the OECD Recommendations on Public Policy Evaluation suggests countries to:

- Establish **institutional mechanisms** to embed evaluation in decision-making processes, both at the organisational level and across government.
- Provide **easy access to evaluations** and present the findings deliberately in order to improve the uptake of evaluation results.

The links between policy evaluations and decision making are still elusive in Portugal. Indeed, there are no instruments in place to understand if and how evaluation results are used. In particular, there are no mechanisms to ensure a connection between *ex post* evaluations and policymaking process or budgetary processes. Evaluation results are rarely discussed at a managerial level within line ministries, presented to high-level meetings or used when discussing budget allocations. For these reasons, the impact of evaluations is still insufficient in the Portuguese system. Moreover, the communication of evaluation results is very limited, hindering the overall impact of the evaluations. Indeed, evaluations remain often unpublished and, even when published, they are often difficult to find as there is no unified database collecting the evaluations.

3.1. Embedding evaluations in decision-making processes

Feedback loops promote the impact of evaluations by creating systematic linkages between evaluation results and decision-making processes. Feedback loops are mechanisms that can be found either at the level of specific institutions, such as management response mechanisms (i.e., when management is required to provide a formal response to the evaluations it has requested), or in cross government decision making, such as through the incorporation of policy evaluation findings into the budget or regulatory cycle or discussions of evidence in strategic planning documents.

To ensure a more systematic impact of evaluations, countries have adopted different types of feedback loop mechanisms, including:

- Use of evaluations for regulatory management, either *ex ante*, through Regulatory Impact Assessments, or *ex post*, through *ex post* regulatory assessments.
- Inserting evaluation clauses in laws, which mandate that an evaluation be conducted after a given period of time (also called sunsetting clauses).
- Including evaluations into budgetary decision making.
- Discussing the results of evaluations at the highest level of government, for instance in the Council of Ministers or Cabinet.
- Management response mechanisms, whereby the commissioning manager commits to providing a formal response to the evaluation's recommendations.

The extent to which these tools are used varies significantly across countries. Only in few countries a majority of these instruments is systematically adopted. A good example are the Netherlands where a good range of feedback loops are in place (see Box 3.1).

Box 3.1. Strengthening the impact of policy evaluation in Netherlands

The Netherlands has put in place concrete mechanisms to ensure the impact of evaluations. The Insight in Quality initiative has further strengthened the link between RIAs and *ex post* evaluations in Netherlands. Indeed, it established the criteria that must be assessed in RIAs (objectives, efficiency and effectiveness) which are the same of *ex post* evaluations (Netherlands Court of Audit, 2021[20]).

In addition, the recent amendment of the Regulation for Periodic Evaluation has required that any policy proposal to the Dutch house of representatives that leads to a substantial policy change (with financial consequences of at least EUR 20 million in a year) must contain an evaluation paragraph, a clause describing how and when the policy proposal will be monitored and evaluated. This is a more concrete measure to increase the impact of evaluations.

Finally, the Netherlands is one of the few OECD countries to have use some management response mechanisms. Indeed, when public evaluations are made available, they are often published with Cabinet responses attached specifying how the Minister intends to deal with the findings and recommendations.

Source: (Netherlands Court of Audit, 2021[20])

In Portugal, *ex post* evaluations are not embedded in specific feedback loops and are hardly discussed in high-level meetings. However, both regulatory impact assessments (RIAs) and *ex post* policy evaluations are co-ordinated by PlanAPP. Indeed, the central body responsible for RIAs is the *Unidade Técnica de Avaliação de Impacto Legislativo (UTAIL)* team and it is now attached to PlanAPP as of early 2022.

Regulatory impact assessments (RIA) can make use of *ex ante* evaluations, as is the case in Portugal. This practice is relatively new in Portugal as it has been implemented firstly in 2017 and was then expanded further in 2018 and 2019 (OECD, 2021[21]). Despite this positive development there are still significant improvements that could be implemented particularly on the engagement of external stakeholders in this process and on the linking of this practice with *ex post* evaluations. E*x ante* evaluations and, more generally, the RIA process, are important for *ex post* policy evaluations as they present an opportunity to clarify a policy's objectives and intended effects, as well as anticipate data collection needs for the *ex post* analysis.

In order to improve feedback loops between evaluation results and decision-making processes, the PlanAPP policy evaluation team would greatly benefit from working with the RIA team within PlanAPP to ensure that the analysis that is conducted *ex ante* for the RIA can improve the quality of *ex post* policy evaluations. To this end, the policy evaluation team could share with UTAIL a methodology to check the evaluability of laws and ask whether the former could share potential data collection needs that arise.

In addition, another way to ensure the impact of policy evaluations is to discuss them at the highest level of government. This is what is done in France, for example, where France Stratégie will share some of its policy evaluations with the Council of Ministers. This is especially the case for the evaluations that are commissioned by the Prime Minister to France Stratégie. In some cases, the evaluations are also discussed in the Council of Ministers.

In order to increase the visibility and impact of its policy evaluations, PlanAPP should consider sharing them with the Portuguese Council of Ministers for information. PlanAPP's position at the centre of government is particularly adequate for this. When developing its multi-annual evaluation agenda, PlanAPP could already discuss with the Presidency of Council of Ministers which ones should be shared with this forum.

Finally, another important way in which governments can ensure that evaluations impact decision making is through the budget process and budgetary decision making more generally. Some countries use evaluations as part of their performance budgeting system. Evaluations can also be useful to inform spending reviews as they can help in understanding which policies are valuable and which ones do not work properly.

In the Portuguese system, spending reviews are still being put in place. In 2016, Portugal has established a working group to study the implementation of a spending review (CFP, 2021[22]). Since then, a minimalist version of expenditure review has been published in a report accompanying the State budget proposal (OPE). However, as analysed by the Conselho das Finanças Públicas, which is the Portuguese Public Finance Council (CFP), there was no published data on the effective realisation of savings under this expenditure review exercise (CFP, 2021[22]).

In order to strengthen future spending review processes and increase the availability of useful evidence to analyse the effectiveness and efficiency of expenditures, PlanAPP could consider, on the longer term, taking into account evidence needs for future spending review rounds in its multi-annual evaluation plan.

3.2. Publicity and communication

Publicity and communication of evaluations are essential to promote their impact. Indeed, policymakers and stakeholders cannot use evidence and the results of evaluation if they do not know about it. For this reason, it is essential to promote the publicity and accessibility of results. However, publicity alone is not enough, and active communication and dissemination strategies are needed to promote impact.

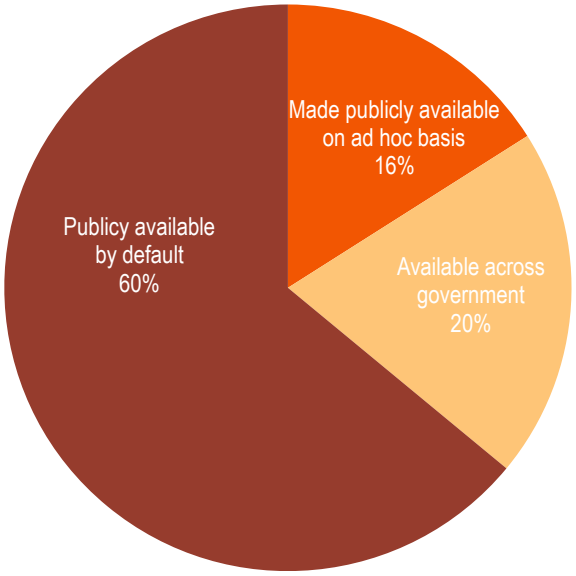
At the moment, in Portugal, with the exception of the evaluations managed by AD&C, the results of evaluation conducted by public administrations are not systematically made public even though they may be made available upon request. This lack of publicity hinders the visibility of evaluations to potential users and thus their impact. Moreover, publicity is important to foster trust and promote a culture of transparency.

The OECD Recommendation states that governments should provide easy access to evaluations and present the findings deliberately to improve the uptake of evaluation results (OECD, 2022[1]). To do so, the Recommendation suggests that countries:

- Make the results of evaluation findings and recommendations **public by default**.
- **Tailor** the way evaluation evidence is presented and communicated to its potential users, in terms of timing, communication channel, format and messaging, by developing a dissemination strategy.
- Make use of **evidence synthesis methodologies** to aggregate evaluation findings and assess them in a systematic manner.

However, publicity of evaluation still varies across OECD countries with 16 OECD countries making policy evaluations available by default. (Figure 3.1)

Figure 3.1. Publicity of policy evaluations



Note: n= 25 OECD countries, option "Other" is not included.
Source: OECD (2020[21]), *Improving Governance with Policy Evaluation: Lessons From Country Experiences,* OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/89b1577d-en.

In addition, publicity of evaluations is not enough to ensure that these will be easily accessed and used. Indeed, in order to achieve impact, several countries have also made use of more active communication strategies. For example, through the creation of online databases or evaluation portals, governments can ensure an easier access to all available evaluations and increase chances of impact (see Box 3.2). Together with this, countries have also worked to make evaluations more friendly to users through creating plain language summaries or providing intuitive graphics. These are all essential efforts to really foster adoption of results.

Box 3.2. Evaluation/ evidence portals: Lithuania and the Netherlands

In **Lithuania**, STRATA has invested in some interesting user-friendly portals to collect analytical work and impact assessments systematically. The first is the **Analytical Information Library (AIB)**, this was developed in collaboration with the Government Chancellery of the Republic of Lithuania and the Lithuanian National Martynas Mažvydas Library. The Library offers a collection of analytical works with more than 2000 evaluations, researchers, and other similar documents. Together with this, a second portal specific on impact assessments has been created containing all *ex ante* analysis conducted from 2020. These are presented as best practices in the field of RIAs as they have all received a feedback from STRATA and respect the established quality requirements

In **the Netherlands**, the Ministry of Finance has a specific website where it is possible to see all evaluations (in particular Policy Reviews and Interdepartmental Policy Investigations) and filter the by department, year, status and so on. This represents a very interesting practice which allow decision makers and citizens to easily access the information.

Source: (Ministry of Finance, n.d.[23]) (STRATA, n.d.[24])

Nevertheless, a first step to increase chances of impact is to make evaluations public. Indeed, evaluations cannot be used if they cannot be accessed by decision makers. A majority of OECD countries publish most of their evaluations, and several of them do so by default – which is to say that evaluations have to be made public unless otherwise stated.

In order to ensure that evidence is available for decision makers to access, the Portuguese government could consider making all evaluations public by default. In particular, PlanAPP's evaluations could be systematically public in order to increase their chances of impact.

In addition to being available, evaluations should be easy to find in order to increase their impact. An important step in ensuring this is to create a unique database where all government evaluations are collected. PlanAPP could consider creating such a database to house all, or most, of the government evaluations. This database could be easy to navigate using keywords and should strive to be exhaustive. Criteria for inclusion in the database would need to be clear and explicit.

Of course, publicity in itself is not enough to ensure the uptake of evaluation results. Results need to be communicated in a way that is understandable and fit-for-purpose for the end user. In Portugal there is a strong need to invest in better communication of results also by creating some user-friendly instruments and tailoring evaluation results to different audiences. This may be explained by the fact that line departments have insufficient time or resources to dedicate to communicating the results of their evaluations.

This is why PlanAPP can play an essential role in supporting the communication efforts of line ministries. To this end, PlanAPP could pair its multi-annual evaluation agenda with a communication strategy, to ensure that along with every planned evaluation, user-friendly deliverables, dissemination events planned and active communication efforts are also conducted to spread results across government. To do so, PlanAPP will need to first identify the relevant users and audiences for each evaluation in the agenda. In addition, on the longer term, PlanAPP could consider amplifying line ministries' communication efforts by reposting some of their social media campaigns and communicating the results of their evaluations on the PlanAPP channels.

Moreover, PlanAPP could strive to position itself as a knowledge broker in the evaluation system. To do so, it could consider synthesising evidence to translate the results from evaluations into information that is easier to understand for decision makers. For instance, evidence gap maps and evidence synthesis methodologies represent interesting areas that PlanAPP could invest in. In particular, the UTA team could particularly benefit from preparing an evidence gap map, which would help with identifying what policy areas require more investment in the future, and thus to inform PlanAPP's multi-annual evaluation agenda.

4 What role for an evaluation champion?

Evaluation champions are essential institutions inside an evaluation system as they considerably help in systematising the production of evaluations, fostering analytical skills across government and setting quality standards for evaluations, as well as promoting the impact of evaluation results.

The importance of evaluation champions is underlined by the OECD Recommendations on Public Policy Evaluation which advises countries to designate evaluation champions to co-ordinate evaluations across institutions and advise on best practices to promote their quality and impact.

Evaluation champions perform different activities depending on their mandate and positioning within government. These can be summarised as follow:

- **Encourage** the conduct and use of evaluations
- **Co-ordinate** evaluations across institutions
- **Advise** on best practices to promote their quality and use, often through manual and guidelines
- **Systemise**: overview of what exists, detecting needs for evaluation, promoting evaluation
- **Support** the technical quality of evaluations by providing peer advice and controlling the quality of the end product of evaluations
- **Promote** impact of evaluations (use, link different evaluations, act as internal knowledge broker)
- **Produce** or **commission** evaluations (but not a key role, see France Stratégie)

As mentioned earlier, several countries rely on one or more institution to co-ordinate policy evaluation across government. However, the exact role, mandate and location within the system vary significantly across countries (see Box 4.1).

Box 4.1. Different evaluation champions in France, Lithuania, The Netherlands and the United Kingdom

France Stratégie - France

The main French evaluation champion today is France Stratégie, a body located within the office of the Prime Minister. It plays a role as a knowledge broker and platform, by collecting and disseminating evaluations from line ministries and other agencies. It is composed of around 40 permanent experts in all of its fields of study and rely on different external scientific advisors. Despite not having a hierarchically superior role vis-à-vis line ministries (as it does not have a legal mandate with sticks and carrots mechanisms), France Stratégie does seek to play an increasing role in promoting evaluations across government. From this position, it performs a dual role:

- performs complex and multi-sectorial *ex post* evaluations directly in co-operation with the line ministries
- supports line ministries in conducting policy evaluations through methodological and quality guidelines.

Evaluation Task Force – United Kingdom

The Evaluation Task Force sits under both the central cabinet office and His Majesty's Treasury. It was established following the 2020 spending review to provide support to HMT spending teams and other departments with their spending decisions, but also received a mandate of helping departments design and deliver high-quality impact evaluations, through the Evaluation and Trial Advice Panel and mobilising the What Works Network (WWN). The overall functions of the new Evaluation Task Force can be divided in three areas:

- scrutinise evaluation plans and business cases submitted to the HM Treasury spending teams
- support and advise departments on evaluations; ensure the use of the results of the evaluations
- promote and encourage the impact of evaluations by making them publicly available.

Ministry of Finance – The Netherlands

In the Netherlands, policy evaluations are co-ordinated by the Ministry of Finance. The Bureau for Strategic Analysis (BSA) is the evaluation champion as it is responsible for overseeing the evaluation policy of all departments. The Dutch approach to championing evaluations differs substantially from that of France Stratégie, insofar as the BSA is much closer to a pure co-ordinating agency. Evaluation functions are organised differently in each ministry, with evaluation departments within ministries responsible for conducting evaluations.

Overall, the BSA functions can be synthesised in:

- co-ordinate and facilitate production of good quality evaluations
- support and advise the ministries with methodological expertise
- oversee the interdepartmental spending reviews.

Lithuania – STRATA

Lithuania's Government Strategic Analysis Centre (STRATA) is the institution endowed with a mandate to improve the country's capacities in regard to strategic foresight, monitoring and evaluation. STRATA is also set to inherit the responsibility to foster the quality and use of *ex post* evaluation, which is conferred to the Ministry of Justice.

Its functions go beyond that of simply co-ordinating policy evaluation and can be summarised as:

- Conducting strategic foresight.
- Supporting line ministries with improving the quality and use of evaluations.
- Conducting evaluations of the National Progress Plan.

In Portugal, PlanAPP was established as an evaluation champion (Presidency of the Council of Ministers, 2021[7]). It was firstly established by the Decree-Law n.21/2021, which attributed to the Centre the functions of planning, design and innovation, *ex ante* and *ex post* legislative impact assessment, monitoring and evaluation of public policies, strengthening engagement with stakeholders , as well as in the training and dissemination of good practices (Presidency of the Council of Ministers, 2021[7]).

The decree n. 646/2022 provides further details on the Evaluation Unit's attributions (PlanAPP, 2022[25]):

- Promote the evaluation of public policies.
- Contribute to the systematisation, development and dissemination of guidelines on the evaluation of public policies (methodologies, procedures and quality control).
- Ensure the process of evaluating and quantifying the economic, social and environmental impact of public policies in conjunction with the services of other government areas.
- Promote collaboration with entities of the scientific and technological system and other entities relevant in the field of public policy evaluation.

The overall mandate of PlanAPP is very ambitious as it covers several areas. However, this is not different from other evaluation champions across OECD countries. For example, STRATA and France Stratégie have very similar mandates to that of PlanAPP regarding policy evaluation, as well as having other important roles in the policy process (foresight and strategic planning). In order to fulfil these functions, it is important that an ambitious mandate is backed up with significant resources, to allow the champion to play its role to its fullest extent. In addition, it can be observed that, in countries where the evaluation culture is strong and well anchored in history (such as the UK for instance), the evaluation champion plays less of a role in conducting evaluations itself. Indeed, as line ministries and agencies conduct credible and robust evaluations themselves, the evaluation champion can focus its role on supporting and advising on the quality and use of evaluations. Therefore, the role of an evaluation champion should be seen as complementing that of line ministries, and something that can evolve over time.

In order to play its role, PlanAPP will need to create a network of evaluators inside line ministries. This network can build on the existing interministerial network for co-operation and sharing of knowledge and resources (*Rede de Serviços de Planeamento e Prospetiva da Administração Pública*, Replan), which is co-ordinated by PlanAPP. The network is still in a developmental phase but was set up with the main objectives of boosting co-operation between different governance areas, promoting capacity building and sharing good practices, as well as harmonising planning procedures and instruments (Presidency of the Council of Ministers, 2021[7]).

In addition, PlanAPP will require a variety of skills and capacities within its evaluation team in order to perform the responsibilities that have been described in this paper. Indeed, as it was shown, evaluation champions are expected to perform a very broad range of activities. These activities require managerial skills (co-ordination and steering functions), communication skills (to provide line ministries with a platform and a communication strategy) and evaluation skills (to support line ministries and also to perform some cross-sectoral evaluations). For this reason, the evaluation unit team will need to undergo specific trainings and attract some more technical figures to support the work of the division.

Summary of Recommendations

Establish a cross-government evaluation agenda to provide high-level guidance on policy evaluation (medium term)

- Develop a multi-annual evaluation agenda in consultation with line ministries, taking into account cross government evidence gaps as well as evidence needs for future Spending Reviews.
- Publish the multi-annual evaluation agenda and share it with the Presidency of the Council for feedback.
- Share the multi-annual evaluation agenda with Parliament and the Court of Accounts for information.

Develop guidelines to clarify the goals and methods of policy evaluation (medium term)

- Develop guidelines on policy evaluation and involve relevant internal and external stakeholders in the creation of these guidelines.
- Include an official definition of policy evaluation in these guidelines.

Foster networks to promote the quality of evaluations and the availability of evaluative evidence for decision making (long term)

- Manage a network of government evaluators to share good practices and methods on policy evaluation. Consider leveraging the existing Replan network to this end.
- Establish an advisory group / peer review group to support line ministries in ensuring the quality of their evaluations.
- Cultivate networks with academia to increase the evidence/ science to policy interface, and thus promote the supply of relevant evidence for policymaking and, conversely, increase the impact of scientific and research evidence.

Strengthen the availability of policy analysis skills in the public sector (long term)

- Develop trainings on policy evaluation, in co-operation with academia, for line ministries.
- Invest in PlanAPP's evaluation team skills, in terms of analysis, as well as in terms of evidence synthesis and commissioning of evaluation.

Ensure access to high quality data to conduct policy evaluations (medium term)

- Obtain recognition as a research institution to facilitate access to data for policy evaluation.
- Develop a framework agreement with the National Statistics Office (INE) and the other national authorities (ONAs) to anticipate PlanAPP's data needs over a multi-annual period (for instance coinciding with the evaluation agenda).

Increase communication activities and make use of evidence synthesis tools to increase the impact of policy evaluations (medium term)

- Discuss and share the evaluations performed by PlanAPP with the Council of the Ministries.
- Use evidence synthesis methods to communicate evaluation evidence effectively to decision makers.
- Create a user-friendly evaluation database to facilitate access to results.

- Develop a communication strategy supporting the dissemination of results from the multi-annual evaluation agenda.
- Organise outreach events in Parliament to present the results of evaluations and foster demand.

Increase the link between ex ante and ex post evaluation to increase their impact (medium term)

- Work between the evaluation and the RIA team in PlanAPP to ensure that Regulatory Impact Assessments can anticipate on data collection needs for *ex post* evaluations and clarify the goals of policies.

In addition, in order to support PlanAPP's efforts to promote the use of policy evaluations, the **Portuguese government** may wish to consider:

Improving the governance of data for evidence-informed policymaking, including by:

- Amending the law on access to microdata for scientific purpose to ensure access to line ministries performing evaluations.
- Developing a data strategy for evidence-informed decision making.

Developing a legal framework providing high-level guidance on policy evaluation, which includes:

- Defining the actors, periodicity and criteria for conducting policy evaluations.
- Making the mandate of line ministry explicit in regard to policy evaluation and clarifying the resources allocated to this activity.

References

Assemblée nationale (2016), *LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique (1) - Légifrance*, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033202746/ (accessed on 26 October 2022). [14]

Baïz, A. (2022), *Quelles évaluations des politiques publiques pour quelles utilisations ?*, France Stratégie, https://www.vie-publique.fr/sites/default/files/rapport/pdf/285590.pdf (accessed on 9 August 2022). [13]

Berg, B. et al. (2020), *Remot Access to Microdata , Final Report*, https://www.cbs.nl/-/media/_pdf/2021/28/final-report-access-to-microdata-annotated.pdf. [16]

CFP (2021), *Economic and Budgetary Outlook 2021-2025*, https://www.cfp.pt/uploads/publicacoes_ficheiros/cfp-rel-07-2021[1].pdf. [22]

Dutch Parliament (2003), *Statistics Netherlands Act*, https://unstats.un.org/unsd/dnss/docViewer.aspx?docID=425#start (accessed on 26 October 2022). [15]

Eurostat (2022), *Recognised research entities*, https://ec.europa.eu/eurostat/documents/203647/771732/Recognised-research-entities.pdf. [12]

France Stratégie (2022), *How can qualitative and quantitative methods be articulated to assess the impact of public policies?*, https://www.strategie.gouv.fr/publications/articuler-methodes-qualitatives-quantitatives-evaluer-limpact-politiques-publiques. [17]

Government of the Republic of Lithuania (2021), *308 On the methodology for the ex-post evaluation of the impact of the existing legal regulation and the format of the report, pa...*, https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/6cb6bfd4b15311eba6328c92adabc234?positionInSearchResults=14&searchModelUUID=b1579e87-5b04-48bf-ab39-11afef4b3b3c (accessed on 19 October 2022). [19]

HM Treasury (2022), *The Green Book, Central Government Guidance on Appraisal and Evaluation*. [4]

INE (2022), *Cooperation Protocol between INE and DGEEC*, https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_cont_inst&INST=387114&xlang=en (accessed on 27 October 2022). [11]

Mathot, A. and F. Giannini (2022), "Evaluation Framework and Practices: A comparative analysis of five OECD countries", *OECD Journal on Budgeting*, https://doi.org/10.1787/911cc792-en. [9]

Ministry of Finance (2022), *Toolbox Policy Evaluations*, https://www.toolboxbeleidsevaluaties.nl/ (accessed on 2 November 2022). [6]

Ministry of Finance (2011), *Guidance on the application of the methodology for evaluating programmes*. [18]

Ministry of Finance (n.d.), *Planned and conducted research*, https://www.rijksfinancien.nl/beleidsevaluatie/onderzoek. [23]

Netherlands Court of Audit (2021), *Insight into Quality – successful operation?*, https://english.rekenkamer.nl/publications/reports/2021/04/21/insight-into-quality (accessed on 14 October 2022). [20]

OECD (2022), "Recommendation of the Council on Public Policy Evaluation, OECD/LEGAL/0478"*, OECD Legal Instruments*, OECD/LEGAL/0478, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0478. [1]

OECD (2021), *Mobilising Evidence at the Centre of Government in Lithuania: Strengthening Decision Making and Policy Evaluation for Long-term Development*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/323e3500-en. [5]

OECD (2021), *Portugal Indicators of Regulatory Policy and Governance*, http://oe.cd/ireg. (accessed on 25 November 2022). [21]

OECD (2020), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/89b1577d-en. [2]

PlanAPP (2022), *Despacho n.646/2022*. [25]

PlanAPP (2022), *Institutionalisation of Policy Evaluation in Portugal: context and challenges*. [3]

Presidency of the Council of Ministers (2021), *"Decreto-Lei n.º 21/2021 de 15 de março", [Law-decree n.21/2021 of the 15th of March]*. [7]

Presidency of the Council of Ministers (2013), *Decree-Law No. 140/2013, D.R. No. 202, Series I 2013-10-18*, https://www.adcoesao.pt/wp-content/uploads/decreto_lei_140_2013_0.pdf. [8]

STRATA (2022), *Training and consultations*, https://strata.gov.lt/poveikio-vertinimas/mokymai-ir-konsultacijos/ (accessed on 15 December 2022). [10]

STRATA (n.d.), *Tools*, https://strata.gov.lt/en/veikla/irankiai/. [24]