



Boosting evidence-based policy making for economic development policies in Italy



DASHBOARD

Evolution	Metric	Actual vs Target	Actual	Target
	Revenue		\$3.4M	\$2.0M
	Profit		\$1.2M	\$0.7M
	Avg. Order Size		\$650.3	\$1.0M
	Online % yearly		99.0%	98.0%
	ROI % sales		15.0%	14.0%
	Cust. Satisfaction		98.0%	95.0%



0001		0.45	▲	+0.45%
0002		-0.23	▼	-2.34%
0003		-1.01	▼	-1.89%
0004		0.02	▲	+0.21%
0005		+2.58	▲	+3.05%
0006		-0.14	▼	-1.42%
0007		-0.73	▼	-0.90%
0008		+1.08	▲	+5.12%
0009		-0.87	▼	-3.88%
0010		-0.65	▼	-1.37%

The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD.

This document was produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

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.



+6.5%

Trends

More info >

CONTENTS

Foreword	3
Acknowledgements	4
Executive Summary	5
1. Main findings and recommendations	7
Summary of main findings	7
Proposed recommendations and actions	9
2. Enabling conditions for evidence-based policy making at MIMIT	12
Overview of the current institutional set-up	12
Overview of existing data sources and data collection and integration capabilities	15
Overview of past and ongoing monitoring and evaluation exercises of MIMIT policies	22
3. Strengthening the Centro Studi's governance drawing on international best practices	31
Setting a strategic framework to achieve Centro Studi's objectives	31
Identifying key stakeholders and their priorities	32
Strengthening capacity for analysis	32
Producing high-quality results	36
Communicating results in an effective way	38
4. Enhancing the Centro Studi's data capacities drawing on international best practices	39
Data collection: administrative and survey data	39
Data integration, harmonisation, and interoperability	42
Improving data capabilities: data infrastructure and skills	44
Data sharing: granting micro-data access for research and building strong collaborations with academia and research centres	46
Developing effective data visualisation tools to support evidence-based policymaking	51
5. Developing the practice of quantifying policy impacts: key methodological considerations and international best practices	52
Identifying the key questions to answer in the impact evaluation	52
Selecting the appropriate methods	54
Overview of the main quantitative and qualitative evaluation methods	55
Quantifying policy impacts: examples from OECD countries	58
References	69

Figures

Figure 2.1. Simplified MIMIT Organisation chart (until March 2022)	12
Figure 2.2. RNA information flow	16
Figure 2.3. Distribution of firms and employment by State aids – 2018	17
Figure 2.4. The process of conducting the Pilot Evaluation	29
Figure 3.1. UK BEIS evaluation governance	37
Figure 4.1. The data value cycle	40
Figure 4.2. Integrated databases in Stats NZ on individuals and business: Integrated Data Infrastructure (IDI) and Longitudinal Business Database (LBD)	43
Figure 4.3. Categories of Cloud service models and implied characteristics	47
Figure 4.4. Different cloud deployment models	47
Figure 4.5. BA/IAB data sources	49
Figure 5.1. Key evaluation criteria	52
Figure 5.2. Impact evaluation along the phases of policy design and implementation	55
Figure 5.3. Visualisation of a simple RCT	60

Tables

Table 1.1. Recommendations and actions for scaling up capabilities and impact of MIMIT's Centro Studi	10
Table 2.1. Research questions for the evaluation of the "Nuova Sabatini"	28
Table 4.1. Data warehouse vis-à-vis data lake: strengths and weaknesses	45
Table 4.2 – In-house servers vis-à-vis Cloud-based solutions: strengths and weaknesses	46
Table 4.3. Mode of micro-data access for research	48
Table 5.1. Overview of evaluation criteria and sample evaluation questions	53
Table 5.2. Overview of evaluation criteria and methods	59
Table 5.3. Case study: evaluating business response rates to redesigned institutional letters in Denmark	60
Table 5.4. Steps of impact estimation for simulation models	61
Table 5.5. Case study: Ex-Ante Evaluation of Investments in Knowledge, Learning and Innovation in Germany	61
Table 5.6. Case study: Ex-ante CBA assessment of an enterprise loan scheme in Ireland	62
Table 5.7. Case study: A systematic review of sustainable practices for businesses	63
Table 5.8. Case study: Evaluating in-itinere Firms' growth patterns in the Netherlands	64
Table 5.9. Case study: evaluating the effect of direct public support for innovation in the UK	65
Table 5.10. Case study: evaluating the R&D Tax Credit in Ireland	66
Table 5.11. Case study: evaluating the impact of R&D subsidies based on the ERDF population-density rule in Finland	67

Table 5.12. Case study: evaluating R&D Subsidies in Italy	68
Table 5.13. Case study: evaluating EcoENERGY for Alternative Fuels Program (ecoEAF) in Canada	68

Boxes

Box 2.1. Characterising firms' use of State aids using the RNA database and firm balance-sheets information	17
Box 2.2. Spain: embedding evaluation and monitoring in legislation	23
Box 2.3. The SME investment incentive scheme "Nuova Sabatini"	27
Box 2.4. Conducting a sample policy evaluation at MIMIT: The Nuova Sabatini Pilot Evaluation	29
Box 3.1. Good practices on evaluation plans	33
Box 3.2. Outsourcing evaluations: the BMWK model	33
Box 3.3. Professionalisation of policy evaluation in Ireland, the United Kingdom and Canada	35
Box 3.4. Trainings to develop policy-makers skills to use evidence, UK and Portugal	35
Box 3.5. Evaluation guidelines in the United Kingdom	36
Box 3.6. Peer Review Group (PRG) in the UK Department of Business and Enterprises (BEIS)	37
Box 3.7. Scientific Council of Growth Analysis in Sweden	37
Box 3.8. UK BEIS's Evaluation Networks	38
Box 4.1. The use of administrative data to support evidence-based policy making	41
Box 4.2. Designing surveys to support evidence-based policy	41
Box 4.3. Integrating around a spine: the experience of Statistics New Zealand	43
Box 4.4. Best practices on data interoperability	44
Box 4.5. Data warehouse and data lake: data management systems in the era of big data	45
Box 4.6. Data storage: In-house servers and Cloud-based solutions	46
Box 4.7. The Research Data Centre at the Institute for Employment Research (RDC-IAB) of the Federal Employment Agency – Germany	49
Box 4.8. Visiting scholars at INPS: a way programme to attract researchers	50
Box 4.9. New Brunswick Institute for Research, Data and Training (NB-IRDT) – Canada	50
Box 4.10. Workforce Success Measures: a dashboard of workforce program participation to support policy-makers – Ohio, US	51
Box 5.1. The main challenges of counterfactual evaluation and ways of addressing them	57
Box 5.2. A novel framework to categorise policy evaluations methods	58

Foreword

OECD countries are placing increasingly greater emphasis on policies aimed at supporting business and economic development. The design and implementation of these policies need to be informed by good evidence on their expected and actual impacts to understand whether support has been effective and ensure that resources and interventions are deployed effectively and sustainably.

Assessing impacts requires good evidence, data, use of the appropriate methods to assess and interpret results, and skills and resources to do so. The Ministry of Enterprises and Made in Italy (MIMIT) has made important efforts to build capacity to supply evidence and evaluation and promote a stronger culture of evidence-based policy making within the institution. The recent establishment of a new Analytical Unit known as Centro Studi responsible for conducting policy evaluations at MIMIT marks an important step towards enhancing the Ministry's capacity to produce and use impact evaluations. This report presents the result of work aimed at supporting the development of the Centro Studi and the use of evidence at MIMIT by providing guidance, practices and methodologies to strengthen the Centro Studi's governance, its data and evaluation capacity and its ability to inform the design and delivery of MIMIT's economic policies in policy domains such as industrial policies, innovation, entrepreneurship and productivity.

Acknowledgments

The work underlying the report is the result of a joint effort bringing together expertise from the OECD Economics Department (ECO), the Science, Technology and Innovation Directorate (STI) and the Public Governance Directorate (GOV), under the guidance of Alvaro Pereira, Director, Country Studies (ECO), Andy Wyckoff, Director (STI), and Elsa Pilichowski, Director (GOV). The work was conducted by a multidisciplinary team composed of Filippo Cavassini, Senior Economist, Fatima Talidi, Policy Analyst, Francesca Papa, Junior Policy Analyst (ECO), Chiara Criscuolo, Head of the Productivity, Innovation and Entrepreneurship Division, Francesco Manaresi, Economist; Ilaria Goretti, Policy Researcher, Antonio Ughi, Junior Policy Analyst (STI), Stéphane Jacobzone, Senior Advisor, Silvia Picalarga, Junior Policy Analyst (GOV). The report received extensive inputs from Isabell Koske, Deputy Director, Country Studies (ECO). Mathieu Miranda and Gemma Martinez (ECO) provided support to the project. Andrew Esson prepared the report for publication.

The project was funded by the European Union via the Technical Support Instrument, and implemented by the OECD, in cooperation with the Directorate-General for Structural Reform Support of the European Commission. Special thanks to Kaspar Richter, Head of Unit, and Stratis Kastrissianakis, Policy Officer, Sustainable Growth and Business Environment Unit.

The work relied on close co-operation with the Italian Ministry of Enterprises and Made in Italy (MIMIT). Special thanks to Federico Eichberg, MIMIT's Chief of Cabinet, Paolo Quercia, Director of MIMIT's Centro Studi, Roberto Sampiero, Director, Division V, Relations with the EU and international affairs; Paolo Casalino, Director General of MIMIT's PNRR unit and former Head of Division, Relations with the EU and international affairs and all the colleagues from across the Ministry who participated in the project. Alberto Cagnazzo and Roberto Spallone (formerly with the Ministry) supported the project in the early phases.

The report benefitted from the inputs and insights from the participants in two Study Visits with OECD peer countries. Special thanks to Denis Martel, Director General, Networks and Collaboration, Strategic Innovation Fund, and Kimberley Accardi, Chief Audit Executive and Director General, Audit and Evaluation Branch Innovation Science and Economic Development (ISED) (formerly Industry Canada); Benjamin Nefussi, Deputy Director, Directorate General for Enterprises, Ministère de l'Économie, des Finances et de la Souveraineté Industrielle et Numérique, France; Sylvie Montoux and Remi Lallement, France Stratégie; Marie Leclair, Head of Repertories, Infrastructures and Structural Statistics Department, National Institute of Statistics and Economic Studies (INSEE), France; Stefan Bender, Head of Division, Research Data and Service Centre, Central Bank of the Federal Republic of Germany (Deutsche Bundesbank); Ruadhán Branigan, Economist, Budget Strategy & Coordination Unit, Department of Public Expenditure and Reform; Dermot Coates, Chief Economist; Cairen Power, Principal, Data and Evaluation Unit; Osama Noujoum, Assistant Principal, Federica De Blasio, Officer, Data and Evaluation Unit, Department of Enterprise, Trade and Employment (DETE), Ireland; Laura Bartiloro, Head of Division, Research Data Centre, Bank of Italy; Paolo Naticchioni, Manager, Research Department, National Social Security Institute (INPS), Italy; Benedikt Vogt, Senior Economist, Department of Labour & Knowledge, Bureau for Economic Policy Analysis, the Netherlands; Rita Bessone Basto, Deputy Director, Guida Nogueira, Research Officer, Office for Strategy and Studies, Ministry of Economy and Maritime Affairs, Portugal; Ignacio Moral Arce, Advisor in evaluation of public policies, Institute of Fiscal Studies, Ministry of Finance, Spain; Rudi Narendran, Senior Analyst and Data Scientist; Ben Walker, Head of Central Monitoring & Evaluation Team, Department for Business, Energy and Industrial Sectors (BEIS), UK.

An early version of the report was discussed at a workshop jointly organised with the delegates of the OECD Economic and Development Review Committee and the OECD Committee on Industry, Innovation and Entrepreneurship. Thanks are extended to the participants for valuable comments and insights.

Executive summary

Evidence-informed policy making is increasingly recognised as a key element for effective public policies. One of the key tools is policy evaluation, which can help governments to design and implement better policies to enhance economic growth and citizen well-being. Enabling better decision-making through ex-ante, in-itinere and ex-post impact evaluation promotes public sector effectiveness and accountability, while offering opportunities for peer learning.

This report aims to support the Italian Ministry of Enterprises and Made in Italy (MIMIT), formerly Ministry of Economic Development (MISE), to strengthen the use of evidence in the design and implementation of policies and interventions to boost economic development. MIMIT is responsible for a broad range of policies, including industrial policies; incentives towards foreign direct investments; internal market, consumer and competition policies; digital agenda and telecommunications. Overall, MIMIT manages, directly or indirectly, a broad set of measures, ranging from incentives, government-backed loans, development contracts and innovation agreements.

The report identifies the enabling conditions for the development and use of evidence to inform policy making at MIMIT and provides a series of recommendations to scale up a newly established analytical unit, known as *Centro Studi*, through improved governance and capacity. Good practices and examples from other OECD countries are provided to help MIMIT enhance skills and capabilities as well as usage of data and quantification. The work underlying this report was carried out within the framework of the Technical Support Instrument of the European Union, in collaboration with the Directorate General for Structural Reform Support of the European Commission.

In recent years, Italy has developed mechanisms to strengthen the use of evidence when developing policies. These include requirements for regulatory impact assessment and a wider statistical system to collect and share data across the public administration, known as SISTAN. Despite recent progress, including MIMIT's initial steps towards setting up an analytical unit, gaps remain, in terms of coordination and information sharing. Across the public administration, often, policy evaluations are performed only as ad-hoc exercises due to limited skills and resources dedicated to these purposes within ministries, and the results are not shared widely.

Despite these challenges, a culture of monitoring and evaluation practices has been emerging within MIMIT, following investments in open data and accessibility. Policy makers at MIMIT are increasingly recognising the value of evaluation. MIMIT collects extensive data on business support measures. Yet, some bottlenecks remain. Data silos sometimes prevent the best use of these resources for policymaking. Linkages to datasets within and outside MIMIT could also be improved to help “connect the dots” and ensure that effects of interventions are measured across sectors and markets. The Ministry's capacity to quantify impacts should also be improved, including by acquiring human and analytical resources to conduct evaluations. Monitoring and evaluation results could also be better embedded in the design of new policies, regulations and economic incentives. Improvements on this front can have important payoffs.

Good practices from peer countries can help equip MIMIT's *Centro Studi* with enhanced capabilities to contribute to policy design and implementation, support collection, integration, and effective use of data and the application of different quantification methods for ex-ante, in-itinere and ex-post evaluation. The work highlights a series of emerging good practices and a commitment to developing further the systematic use of data and analysis in the design and implementation of policies and instruments under the responsibilities of MIMIT.

The proposed recommendations are intended as a comprehensive package focusing on both targeted measures to improve the analytical capacity and governance of the *Centro Studi* as well as more systemic efforts to ensure its ultimate impact on policymaking. These are intended as guidance that will need to be translated into a strategic plan by the *Centro Studi*. Some recommendations could generate results relatively quickly, while others will take more time to be implemented.

For instance, one priority for the short term could focus on recruiting staff with relevant analytical and quantification skills or implementing measures to attract external researchers at the institution. Such mechanisms can be set up relatively quickly, even if building up in-house capacity will take more time. Another short-term area for attention would be the development of the IT infrastructure: acquiring tools and hardware to store data and perform the analyses. This would also increase the relevance of the Centro Studi to support implementation of the NRRP, which would help support strategic core activities of the Ministry and show its added value on quantitative matters.

In parallel, longer-term mechanisms can be envisaged to strengthen the Centro Studi's capacity for supplying high-quality evaluations, such as establishing peer review mechanisms or a Scientific Advisory Council or setting up secondment agreements with relevant institutions (for instance, the National Statistical Institute, the Bank of Italy, the National Social Security Institute). In addition, MIMIT could initiate efforts to develop a comprehensive data infrastructure to be used for policy analysis, by designing and implementing a centralised data lake and developing routines to integrate available data and create ready-to-use integrated datasets for analytical purposes. Once the first evaluations have been conducted, it will also be important to devise a targeted communication strategy, tailoring the communication of evaluation results to meet the needs of diverse audiences, and promoting the use of the results of evaluations in policy and budgetary decisions.



1. Main findings and recommendations

This report proposes a series of policy actions to build up the capabilities, skills and evaluation agenda of the newly established Centro Studi of the Ministry of Enterprises and Made in Italy (MIMIT) that transformed the existing statistical unit into an analytical unit with the mandate of conducting policy evaluation. The recommendations build on an in-depth assessment of the enabling conditions and institutions for using evidence to inform policy making at MIMIT, presented in Chapter 2. The assessment draws on exchanges with MIMIT stakeholders (Directorate Generals for Industrial Policy and SMEs; Business Incentives; Internal Market and Competition; Intellectual Property; Digital Infrastructures, Minister's Cabinet and the General Secretariat), as well as the Bank of Italy, the National Statistical Institute (ISTAT), Invitalia, National Social Security Institute (INPS) and the analytical units of Italy's public investment bank (Cassa depositi e prestiti) and Italy's Employers Federation (Confindustria). The recommendations also build on good practices and guidance from OECD countries, presented in Chapters 3, 4 and 5, including findings of a series of knowledge exchange webinars with relevant bodies in Canada, France, Germany, Ireland, the Netherlands, Portugal, Spain and United Kingdom.

SUMMARY OF MAIN FINDINGS

Some important enabling conditions are in place to step up the use of evidence in policy making

There is a strong case to boost evidence-based policy making in Italy. There have been significant improvements in the last few years, including the systematic adoption of regulatory impact assessment (AIR in Italian). However, gaps remain. There is not a unique entity responsible for coordination across Ministries and policy evaluations are often performed only as ad-hoc exercises.

Despite these challenges, a culture of monitoring and evaluation has been emerging across MIMIT, following investments in open data and accessibility. Policy makers at MIMIT are increasingly recognising the value of evaluation, as evidenced by the creation of the new Centro Studi. While additional capacity and expertise will need to be developed, MIMIT's commitment is a promising precondition for the systematic use of evidence in designing, monitoring and evaluating policies.

Some key enabling conditions are also in place. MIMIT already collects and has access to a wide range of data on the policies under its responsibilities, including on fiscal incentives, beneficiaries and patents. Some good practices in data collection and data presentation already exist, for instance through the National Registry of Incentives (RNA in Italian) and a new database ("Nuova Banca Dati Agevolazioni" or NBDA) that integrates the information collected in the RNA with data on aid payments. Requirements for monitoring and evaluation have also been included in the design of some policies, providing an incentive to use evidence when evaluating policy impacts more systematically. For instance, the legislation informing the "Start-up Act", a package of measures supporting start-ups and innovative firms, includes requirements to collect data on implementation progress and impacts, allocates some resources for this purpose and introduces an obligation to monitor and report back to Parliament on impacts.

These enabling conditions have facilitated the development of evaluations quantifying policy impacts. While these evaluations have been mostly conducted by institutions outside the Ministry, MIMIT has used some of the findings from these evaluations to adjust some policies, for example, by modifying the targeting of the *Fondo di Garanzia*, a public guarantee scheme for small and medium sized enterprises.

Governance, coordination and skills need to be strengthened to take full advantage of these enabling conditions

The decision to establish the Centro Studi has been an important step towards using evidence more systematically for policy making. However, the formal establishment of the unit needs to be supported by a systematic approach to collect data and conduct policy evaluations, together with adequate investment in strong evaluation capacities.

Planning and coordinating policy evaluations needs strengthening. MIMIT is not yet fully “connecting the dots” across the evaluations conducted within and outside the Ministry. For instance, business incentives are being evaluated at the request of individual Directorate Generals (DGs). It will be important to have a more co-ordinated approach to evaluating MIMIT policies to also seize synergies and complementarities. Coordination and networking should also be extended to data collection and usage. Only limited firm-level information on State aids is available in the RNA and the NBDA databases, while other relevant data is kept within DGs, including data on other types of incentives. Information on the data contained in these “data silos” is currently scarce and not structured. Developing standardised procedures to collect data across DGs and stronger incentives for co-operation across DGs on data integration could help reduce data fragmentation. The Centro Studi’s location within the Secretariat General offers an opportunity to have a “bird eye view” on these processes, facilitating integration and coordination while also serving as a conduit for coordinated evidence and analysis to the Ministry’s leadership.

Resources should be aligned with the commendable commitment of MIMIT to strengthening evidence-based policy making. There is a network of statisticians and analysts across MIMIT that can be potentially leveraged. At the same time, it is important for the Centro Studi to invest in staff with statistical, economic and data analysis expertise. Developing partnerships with analytical institutions within and outside the Italian government, for instance by developing secondments and research programmes can also help. It will be equally important to acquire appropriate software, notably statistical packages, and hardware to perform analysis and store data. Acquiring statistical packages is relatively inexpensive compared to the potential benefits of assessing and improving policies that have significant budgetary and growth implications for the Italian economy.

Data should be better integrated and become more accessible to fully exploit the evidence already collected

Attention is also needed towards integrating data from providers outside MIMIT such as the National Statistical Institute (ISTAT), the National Social Security Institute (INPS), the system of the Italian Chambers of Commerce (Unioncamere). Access to these data by the Ministry is currently limited and usually regulated by ad-hoc agreements. In principle the Ministry may access most of the external data through the National Statistical System (SISTAN), but in some cases this access can be limited (for instance, currently ISTAT does not provide data in which firms are properly de-identified, thus limiting the possibility to integrate it with other data sources).

The rich data collected by MIMIT should also be made accessible outside, preserving confidentiality concerns, to facilitate research and evaluations conducted outside MIMIT, from which policy makers can greatly benefit. The Ministry currently lacks a structured and well-designed programme to provide micro-data access to external researchers, relying on ad-hoc agreements, internships or external collaborations. The work of external economists is also hampered by the lack of information on database structure and location, which may slow down research activities.

The integration of a large amount of data, its storage (e.g., in a data lake or data warehouse), and its use for reporting and policy evaluation requires either the presence of a physical IT infrastructure or the use of a cloud solution. Currently, the Centro Studi does not have an IT infrastructure of any of the two types. More generally, the DGs have different IT solutions for data storage and management, thus facilitating the emergence of data silos – which preventing the flow of information and a more effective use of data inside the organisation.

Quantifying policy impacts throughout the policy cycle can strengthen the Ministry’s support for the economy

Tailored guidance on how to quantify policy impacts would facilitate the practice of evaluating policies. Currently, MIMIT has access to general guidance and methodologies, such as the guidelines for regulatory impact assessment and ex-post evaluation by the Department for Legal and Legislative Affairs of the Presidency of the Council of Ministers (Presidenza del Consiglio dei Ministri, 2018^[1]). However, the guidelines focus on the procedural steps and do not provide an overview of quantitative and qualitative methods. For EU-funded programmes, MIMIT has been relying on the comprehensive guidelines developed by the European Commission, which are very exhaustive and of general application (EC, 2013^[2]).

The Ministry could benefit from a better understanding of the full spectrum of quantification methods and when to apply them. With the exception of EU-funded programmes, no ex-ante evaluations are conducted and this is an area for further development. At the same time, not all policies and interventions will need to go through an in-depth evaluation or be evaluated using counterfactual methods. Accordingly, being selective and strategic in what and how is evaluated and what methods can be used would help MIMIT better use resources and improve the use of evidence. Over time, a commitment to continuously improve the quantification methodologies and publish all evaluations will strengthen the evaluations produced by the Centro Studi.

PROPOSED RECOMMENDATIONS AND ACTIONS

Building on these findings, MIMIT could consider a number of actions to ensure that the Centro Studi and the other structures of the Ministry effectively contribute to an enhanced use of evidence at MIMIT. Proposed recommendations are broken down in discrete actions to facilitate implementation. They are presented in Table 1.1. Chapters 3, 4 and 5 offer good practice examples and guidance on which MIMIT can build to inform the implementation of the proposed recommendations and actions.



TABLE 1.1. Recommendations and actions for scaling up capabilities and impact of MIMIT's Centro Studi

Recommendations	Actions
IMPROVING THE GOVERNANCE OF THE CENTRO STUDI	
1. Build internal coordination mechanisms with the Minister, the cabinet and the DGs, including the new NRRP Unit	<ul style="list-style-type: none"> • Conduct regular meetings with internal stakeholders • Set up focal points in each DG • Based on the evaluation agenda, define a multidisciplinary team (<i>tavolo di lavoro</i>) for each strategic evaluation • Establish a network to promote knowledge exchange • Periodically organise seminars to present internal work
2. Build regular partnerships with external stakeholders	<ul style="list-style-type: none"> • Identify key stakeholders and map existing partnerships • Develop focused partnerships (<i>accordi quadro</i>) on areas such as data sharing, secondment opportunities, network of evaluators, and potential joint evaluations of policies
BREAKING DOWN THE DATA SILOS WITHIN THE MINISTRY AND WITH MANAGING ENTITIES	
3. Conduct a census of existing datasets across the DGs and managing entities	<ul style="list-style-type: none"> • Prepare a template metadata report to be provided to DGs and managing entities • Collect and continuously update metadata
4. Set out regulations and procedures to share the data between the Centro Studi and DGs	<ul style="list-style-type: none"> • Assess the current regulatory framework and procedures through which Centro Studi may access data from DGs and integrate them with other data sources, while ensuring confidentiality • Strengthen procedures and regulations to lower the barrier to data access and integration within the Ministry, introducing master data to facilitate the interoperability of the various data sources
STRENGTHENING CAPACITY FOR SUPPLYING EVALUATIONS BY INVESTING IN NEW STAFF	
5. Recruit and retain staff with a range of analytical and quantification skills and expertise in economics, social and behavioural sciences	<ul style="list-style-type: none"> • Map the capacity needs and the skill gaps • Collaborate with the Department of Public Function (Dipartimento Funzione Pubblica) to mobilise skills • Support the HR department on the creation of more tailored staffing requests
6. Encourage secondments with relevant institutions (e.g., ISTAT, Bank of Italy, INPS)	<ul style="list-style-type: none"> • Elaborate the secondment notice based on the partnership agreements signed with relevant institutions. • Promote secondment opportunities and publish them on the InPA website
7. Create programmes to attract researchers on a time-bound basis to increase the supply of evaluations	<ul style="list-style-type: none"> • Contact INPS and potential partner universities • Elaborate a framework programme to attract external researchers • Define a list of micro-data to be shared with externals and set up the data-sharing facility
8. Promote trainings opportunities	<ul style="list-style-type: none"> • Perform stock-taking of existing training programmes within MIMIT (HR Directorate) and outside (e.g., Dipartimento Funzione Pubblica; SNA) • Promote the registration to specific courses both inside the Centro Studi and in all DGs
DEVELOPING THE PRACTICE OF QUANTIFYING POLICY IMPACTS	
9. Develop an evaluation plan	<ul style="list-style-type: none"> • Following interviews with internal and external stakeholders, take stock of potential evaluations to be conducted on key forthcoming and on-going policies • Prepare a multiannual evaluation agenda prioritising a list of evaluations to be conducted and indicating relevant databases, methodologies, timelines and workplan
10. Embed data collection in policy design	<ul style="list-style-type: none"> • Identify and prioritise data needs • Design a plan to systematically ensure that data collection is embedded in policy design
11. Develop descriptive statistics and indicators to help strengthen monitoring	<ul style="list-style-type: none"> • Review existing monitoring exercises and identify areas where monitoring is not currently conducted but could be introduced • Conduct more advanced descriptive statistics to support in-itinere and ex-post impact assessment
STRENGTHENING CAPACITY FOR SUPPLYING EVALUATIONS BY INVESTING IN NEW STAFF	
12. Strengthen ex-ante assessments to start developing quantification practices	<ul style="list-style-type: none"> • Develop standard methodologies for ex-ante assessments and, more specifically, cost-benefit analysis • Identify opportunities to start applying ex-ante assessments to new policies introduced by MIMIT • Apply different ex-ante methodologies in accordance with evaluation questions to be answered and the evaluation plan

TABLE 1.1. Recommendations and actions for scaling up capabilities and impact of MIMIT's Centro Studi (continued)

Recommendations	Actions
13. Plan and start designing counterfactual evaluations for ex-post and in-itinere policy assessments	<ul style="list-style-type: none"> • Scope relevant databases for in-itinere and ex-post assessments • Build on materials from the pilot evaluation of the Nuova Sabatini (Output 5) and the technical report on quantification methods (Output 3) to apply different in-itinere and ex-post methodologies
PROMOTE HIGH-QUALITY TRUSTWORTHY EVALUATIONS	
14. Develop a protocol for implementing evaluations	<ul style="list-style-type: none"> • Develop a taxonomy for applying analytical methods to policy making • Adopt internal MIMIT-specific guidelines for impact evaluation at MIMIT to promote high-quality and trustworthy evaluations and operationalise the guidelines by developing a protocol for conducting evaluations
15. Establish a peer review mechanism	<ul style="list-style-type: none"> • Identify at least two peer reviewers for each evaluation team • Systematise the use of peer reviewers throughout the evaluation process
16. Establish a Scientific Advisory Council	<ul style="list-style-type: none"> • Map scientific councils across similar line ministries in Italy and in the EU • Establish the MIMIT Scientific Council with a number of senior experts
DEVELOPING A COMPREHENSIVE DATA INFRASTRUCTURE TO BE USED FOR POLICY ANALYSIS	
17. Develop the IT infrastructure: acquiring tools and hardware to store data and perform the analyses	<ul style="list-style-type: none"> • Conduct an internal assessment of existing IT systems used within the Ministry • Develop and upgrade the IT systems to provide higher computational power and storage capacity for analytical tasks. • Acquire software to support Centro Studi's analytical tasks
18. Develop routines to integrate available internal and external data and create ready-to-use datasets for analytical purposes	<ul style="list-style-type: none"> • Develop data-setting and data-cleaning routines and codes to automate the integration and processing of data collected in the data lake • Develop routines and codes to extract and provide ready-to-use datasets for analytical purposes
19. Design and implement a centralised data lake	<ul style="list-style-type: none"> • Conduct an internal assessment of existing data management systems used within the Ministry • Design and implement a centralised, coherent, data lake solution
STEPPING UP THE INFORMATIONAL CAPABILITIES OF FIRM SURVEYS	
20. Create a survey registry	<ul style="list-style-type: none"> • Set up, implement and administer the registry database
21. Define guidelines for conducting firm surveys related to the policies of the Ministry	<ul style="list-style-type: none"> • Identify how the survey will complement existing administrative data to answer policy relevant questions • Develop a sampling strategy for beneficiary and non-beneficiary firms • Assess the possibility to perform additional waves of the surveys in the future • Define a series of good practices to identify the external contractors that should administer the surveys
22. Develop a periodic survey to monitor the take-up and usage of the Ministry policies	<ul style="list-style-type: none"> • Plan and administer a periodic survey to monitor the Ministry's policies over time
ENSURING POLICY IMPACT AND DISSEMINATION	
23. Use results of evaluations in policy and budgetary decisions	<ul style="list-style-type: none"> • Discuss with the Legislative office and the Cabinet the possibility to insert evaluation provisions in strategic new policies • Set a meeting with the Cabinet and the Budgetary office to explore the use of evaluation results for significant policy changes
24. Make evaluation results public by default	<ul style="list-style-type: none"> • Develop the potential website page of the Centro Studi, together with the informatic office of MIMIT
25. Tailor communication of evaluation results to users	<ul style="list-style-type: none"> • Identify relevant users and audiences of the Centro Studi. • Develop and introduce a communication strategy, which accounts for the need to tailor communication products • Conduct trainings to enhance the ability of staff to produce communications and notably digital communication
26. Develop a data-visualisation tool to provide key policy-relevant indicators on policy take-up and usage	<ul style="list-style-type: none"> • Choose a set of highly relevant policies by the Ministry to be included in the visualisation tool • Develop a dashboard with selected indicators to be visualised

2. Enabling conditions for evidence-based policy making at MIMIT

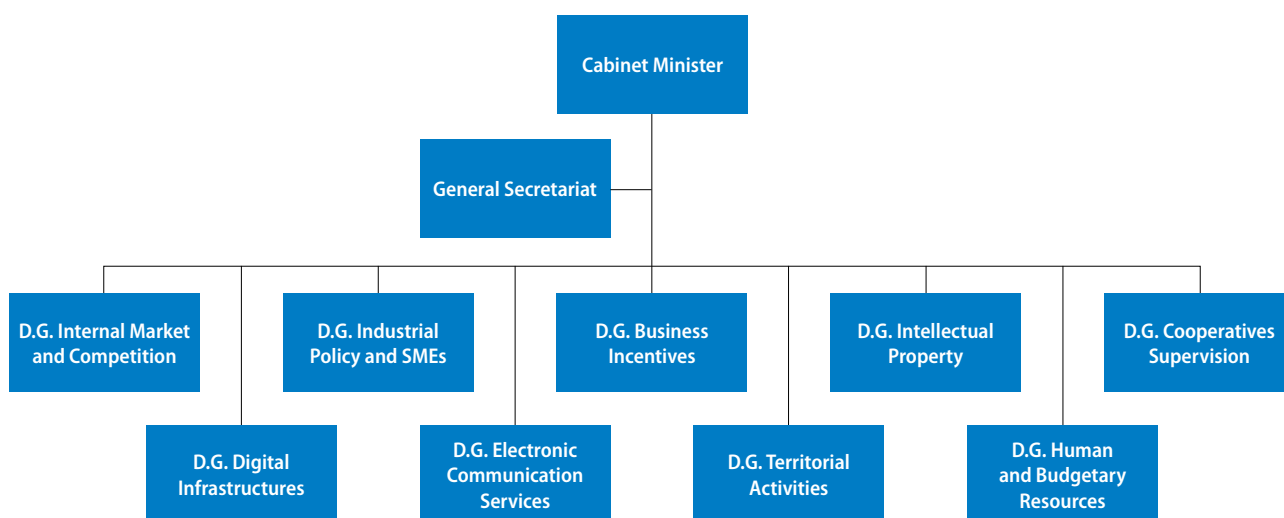
This chapter offers an overview of MIMIT’s current set-up for the monitoring and evaluation of its policies. This includes a discussion of the state of play in the establishment of MIMIT Centro Studi; an overview of existing databases of relevance for the evaluation of MIMIT policies and a review of relevant policy evaluations that have already been conducted.

OVERVIEW OF THE CURRENT INSTITUTIONAL SET-UP

A historically dispersed evaluation system: the role of the D.G.s in performing analysis and monitoring

MIMIT has recently undergone significant structural changes.¹ It is currently composed of nine directorates coordinated by a General Secretariat (See Figure 2.1).

FIGURE 2.1. Simplified MIMIT Organisation chart (until March 2022)



Note: This is a simplified organisation chart. A detailed organigramme is available on the Ministry’s website: <https://www.mimit.gov.it/it/ministero>.

Source: GENERAL SECRETARIAT Division V – Relations with the European Union and international affairs.

Up to the recent creation of the Centro Studi, MIMIT did not have a central body responsible for coordinating the different D.G.s in the evaluation activities. The organisation of monitoring and analytical activities inside MIMIT has historically been decentralised across different D.G.s., with some DGs already performing some level of research or data collection activities related to monitoring and evaluations (see Section 2.1.3).

However, while the Ministry manages a significant amount of data, statistical inference and analysis could be performed in a more systematic way. The work is conducted mainly for administrative monitoring purposes. At the moment, each DG collects separate databases with dedicated statisticians. Data collection and analysis functions were identified in all the five D.G. interviewed as part of the project, with around 15 officials with statistical background. However, there is also heterogeneity across D.G.s in their ability to perform analyses.

1. The latter occurred in 2019 when two DGs were transferred to MAECI (Trade Promotion & International Trade Policy). As part of the process of institutional reorganisation at MIMIT, a new DG on Industrial Reconversion has been set up, while the DG on Territorial Activities is no longer active. In November 2022, the Ministry has adopted its current denomination.

The new Centro Studi: functions and set-up

In 2021, MIMIT decided to increase its capacity for policy evaluation by establishing a new Centro Studi, to replace the former Statistical Office, which was responsible for collecting data and feeding them into the National Statistical System (SISTAN).

The establishment of the Centro Studi fulfils a commitment to strengthen evidence-based policy making expressed in a number of normative acts. These include the Act of Address (*Atto di Indirizzo Politico*) by the Minister, stressing the importance of monitoring activities and policy evaluation,² a Decree of the President of Council of Ministers, mandating the realisation of studies on industrial policies to support digital transformation under the coordination of the General Secretariat, and a provision in the National Recovery and Resilience Plan (NRRP) focusing specifically on the *in itinere* and *ex-post* evaluation of NRRP policies (Art. 7,c.2, DL n.77/2021). The establishment of the Centro Studi is in line with similar processes in other Italian Ministries, such as the Ministry of Economics and Finance (MEF)³ or the Ministry of Justice⁴ that have set up their own analytical units.

The responsibilities and functions of the MIMIT Centro Studi include:

- Coordination of the statistical activities carried out by MIMIT's Directorate Generals;
- Collection, processing, analysis and dissemination of statistical data in collaboration with ISTAT;
- Providing statistics and information for the needs of other ministries and public institutions;
- Evaluation of projects included in the NRRP;
- Ex-ante, in-itinere and ex-post evaluation and impact assessment of MIMIT policies in conjunction with the relevant Directorate Generals.

The detailed functions of the unit are further specified in the Decree of the Minister 19 November 2021.⁵

While the old Statistical Office was temporarily located, until March 2022, within Division V (Relations with the European Union and international affairs) of the General Secretariat,⁶ the Centro Studi will operate as a separate division (i.e., Division IV) of the General Secretariat, and will supervise impact evaluation activities across Directorate Generals.

This change from being located inside one D.G. to the General Secretariat will allow the Unit to take a cross-cutting view on MIMIT work. At the same time, it is also getting closer to the decision makers of the Ministry, which might allow the Unit to be better mobilised to support policymaking. Still, this will require strong capacity and professional skills with solid and impactful evaluation results.

As of May 2023, the Centro Studi is undergoing a process of restructuring and development. A new Director of the Division has been appointed and the process of recruiting additional staff is ongoing. Three new statisticians were hired with long-term contracts, bringing the total number of the Centro Studi's staff to 5 (four statisticians and one market analyst). Three new staff (economists and experts of industrial relations) will be hired in the near future, with short-term contracts up to two years. Two experts of evaluation of public policies are expected to join the Centro Studi in the second half of 2023, bringing to total number of staff to 10.

2. Priority II in the *Atto di Indirizzo Politico* deals with rationalising the scheme of incentives and enhancing intervention in favour of SMEs and includes an element of monitoring activities and policy evaluation

3. For more details on MEF's analytical unit (*Unità di missione Analisi e Valutazione della Spesa*), see: https://www.rgs.mef.gov.it/VERSIONE-I/ragioneria_generale/struttura_e_funzioni/ispettorati_generali_e_servizio_studi/unita_di_missione_avs/

4. For more details on Ministry of Justice's analytical unit (Direzione generale di statistica e analisi organizzativa), see: https://www.giustizia.it/giustizia/it/mg_12_2_6.page

5. As detailed in Decree of the President of the Council of Ministers, July 29 2021, n.149, footnote 2, the detailed functions of the Centro Studi include: Performing study and research activities in subjects of strategic interest to the Ministry; Supporting and coordinating activities for statistical analysis within the competence of the Ministry, in connection with the National Statistical System (SISTAN), and the National Institute statistics (ISTAT); Carrying out research and analysis to increase effectiveness and the efficiency of administrative action; Supporting study research and investigation in conjunction with the Directorates-General competent in the field of statistics; Promoting statistical surveys.

6. The former "Statistical Office" was initially located inside the D.G. for energy policies and used to have a significant focus on energy data. With the transition of energy competences to the newly created Ministry of Ecological Transition (MiTE), the "Statistical Office" relocated with the remaining functions inside the General Secretariat in Division V.

MIMIT's relationships with key stakeholders

MIMIT has established several collaborations with external stakeholders. External stakeholders have supplied complementary datasets and data expertise to MIMIT (e.g., Istat, Camere di Commercio). In other cases, these stakeholders have often helped MIMIT to strengthen analysis and evaluations (e.g., Invitalia, Universities). The existence of a network represents an important asset to expand the capabilities of the Ministry. However, at the moment, these relationships would require further systematisation.

Key stakeholders include:

- 1. Invitalia.** Invitalia is an agency under the Ministry of Economy and subject to “*vigilanza*” by MIMIT. It is responsible for all the incentives on new enterprises and on innovative start-ups, and also offers technical support through the creation and management of databases and ICT infrastructures. In particular, for MIMIT, Invitalia is responsible, for the management and monitoring of policies financed with the EU Cohesion Fund, such as the *PON Imprese e Competitività* and *PON Iniziativa PMI*. Nevertheless, Invitalia mainly performs qualitative and quantitative analysis based on surveys and focus groups, rather than counterfactual analysis. With its expertise in monitoring, secondments from Invitalia could be useful for the Centro Studi to draw on additional, complementary expertise.
- 2. ISTAT (SISTAN).** The National Statistical Office (Istat) represents an important partner for MIMIT, as it not only collects vast amount of relevant data, but also already performs some evaluations in the field of industrial policies and business incentives, such as the *Dossier Analisi dei provvedimenti fiscali sulle imprese*. The Centro Studi is MIMIT's contact point inside the National Statistical System (SISTAN), i.e., a large network of national institutions and public administration coordinated by Istat which collects data for statistical purposes. Being part of the SISTAN allows MIMIT to access all Istat microdata (e.g., data on Italian firms, advanced estimation models) through a simple request. Nonetheless, for privacy concerns, the potential of these datasets is yet not fully exploited (see Section 2.1.2). Through SISTAN, the Centro Studi will participate to the “*Circoli di Qualità*” and to the *Programma Statistico Nazionale*, in coordination with Division XII of DG Market (DGMCCVNT).

To enrich the quality of its analyses, the Centro Studi could benefit from a collaboration with Istat. In the past, this has happened on ad-hoc basis through the use of a *protocollo d'intesa*, despite some attempts to establish a more systematic collaboration in 2016. To make the collaboration between these two institutions more systematic, the two institutions could arrange an *Accordo quadro*, which offers a stronger framework. Istat has already a systematic collaboration of this kind with Bank of Italy. If legally feasible, it would even be preferable to organise a tripartite framework agreement among MIMIT, ISTAT and the Bank of Italy.

- 3. Bank of Italy.** The Bank of Italy represents a stronghold in Italy with respect to analysis and evaluation. Bank of Italy has often accessed MIMIT's databases to perform analysis and evaluations, such as for studies on business incentives for innovative start-ups. However, the collaboration between MIMIT and the Bank has always been ad-hoc. Despite this, establishing a more stable collaboration is possible through the creation of a framework agreement. In the context of the NRRP, Bank of Italy collaborates with MIMIT and MEF for the evaluation of Industry 4.0. This could represent an important first step in setting the basis for more structured collaborations on evaluation. The strength of the Bank of Italy with regard to evaluations and economic quantification is that it is equipped with very skilled evaluators, economists, statisticians and data scientists. In this respect, the Bank of Italy represents a key asset that could be mobilised by the Centro Studi in the future.
- 4. Cassa depositi e prestiti – Centro studi.** Cassa depositi e prestiti (Cdp) is one of the main actors financing private and public investments in Italy. Cdp has a dedicated Centro Studi that for several years has produced high quality analysis and macroeconomic reports. Only recently, following a restructuring of functions, the Centro Studi became responsible for monitoring and impact evaluations, with a new team of 8 highly trained economists. The advantages of Cdp in setting up a Centro Studi are the already strong analytical capacities inside the organisation, its good reputation and consequently its ability to attract interesting profiles. Despite this, it will need to invest on data

access as it is not part of the SISTAN. Being at an initial stage of their evaluation activity, they can represent an interesting counterpart MIMIT Centro Studi to share experiences.

- 5. Confindustria – Centro studi.** As a business association, Confindustria is another interesting partner for MIMIT, interested in most policy domains covered by the ministry. Confindustria itself is well equipped with a Centro Studi which has the objective of providing evidence and knowledge to its members. The way this Centro Studi functions could provide interesting insights for the setting up of MIMIT Centro Studi. The Confindustria Centro Studi was formally established in 1970 to provide evidence, technical knowledge and to facilitate innovation both for the Italian business association and for the Italian economic system more broadly. Today it focuses on four main thematic areas: 1) Economic analysis and forecasts; 2) Geo-economic scenarios; 3) Trends in business and industrial systems and 4) Evaluation of public policies. While the department is relatively small (≈ 17 people), it is equipped with skilled economists who are able to design and perform econometric analysis. When additional skills are needed (engineering, legal or fiscal) the Centro Studi has also efficient connections with Confindustria's policy departments where these skills are more common. The Centro Studi also collaborates with external stakeholders like ISTAT, MEF, MIMIT and Bank of Italy. They represent important sources of data for their analysis.
- 6. Nucleo di valutazione e verifica degli investimenti pubblici (NUVAP).** NUVAP, an office located inside the Legislative Department of the Presidency of the Council, offers technical, methodological, support to all public administration on the monitoring and evaluation of public investments in the area of infrastructures, research, environment, innovation. It provides MIMIT with orientations and support on planning and implementing inquiries, as well as on developing programme indicators and incentive mechanisms. MIMIT is also actively involved in working groups, and technical seminars developed in the context of the National System for the Evaluation of Cohesion Policies (*Sistema Nazionale di Valutazione delle politiche di coesione*) coordinated by the NUVAP. This engagement enables the Ministry to collect further information on evaluation techniques, statistical analysis methodologies related to the territorial development policies.
- 7. INPS.** The Italian social security institute (INPS) is a particularly useful reference as an example of an institution that was able to attract skilled labour force with the creation of dedicated programmes to attract high level academic experts. In particular, the experience of the VisitINPS program, which benefitted from strong leadership in recent years, shows the potential for change in the Italian administration with additional leadership and a reinforcement of capacity (See Box 4.8). This also offers a good example of partnerships with outside researchers and academics, which can inform the analytical work of public administrations.

OVERVIEW OF EXISTING DATA SOURCES AND DATA COLLECTION AND INTEGRATION CAPABILITIES

Available MIMIT datasets to monitor incentives and State aids

The Ministry regularly receives information on the beneficiaries of incentives and State aids, either by collecting the data directly or by obtaining them from other entities that are in charge of managing the measures.

The main databases that collect State aids information are the National Register of State Aids (Registro Nazionale degli Aiuti, RNA) and the “Nuova Banca Dati Agevolazioni” (NBDA). In addition to those, the Ministry collects datasets on specific MIMIT incentives and State aids that are accessible to personnel of the offices managing those regimes.

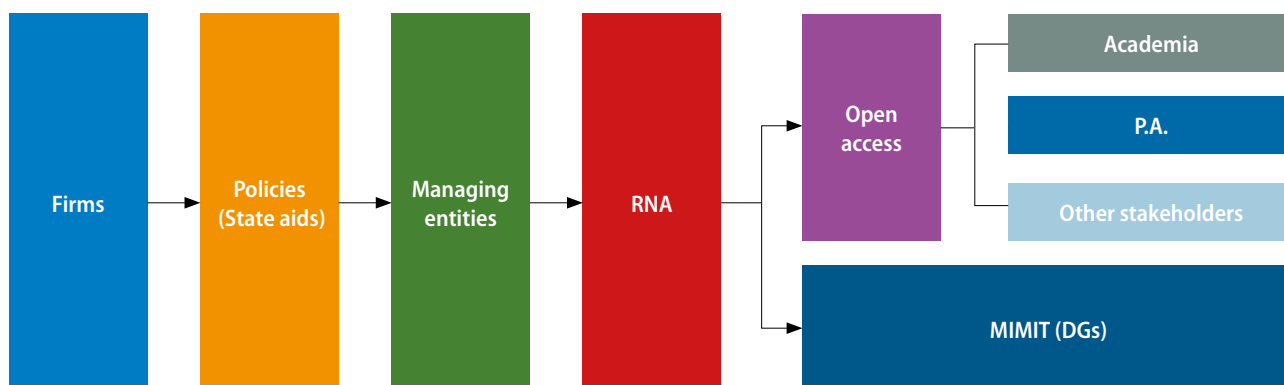
National Register of State Aids (RNA)

Since August 2017, MIMIT's DG for business incentives (DGIAI) manages the RNA database, a centralised database that collects monthly firm-level information on all State aids in Italy. Previous to the RNA, information on beneficiaries of aids was not standardised and highly fragmented across managing entities.

The RNA has been created for legal-administrative purpose, to verify that State aids are granted in compliance with European legislation. Figure 2.2 shows the inflow of information that supports the RNA database, and the outflow

to users. When a firm requests a State aid, the managing entity (e.g. a Ministry or another public or private institution acting on its behalf) uses the RNA to check that the cumulated amount of aids granted to the firm does not exceed the ceilings set by EU regulations. If this and other regulatory checks are passed, the aid is granted, and the managing entity is legally required to register it in the RNA. Information collected in the RNA includes: the name and fiscal code of the beneficiary, its region and sector of activity, the State aid granted and its amount, and the managing entity.

FIGURE 2.2. RNA information flow



Note: the figure shows the flow of information leading to RNA updates and its accessibility to users.

Source: OECD elaboration.

The RNA database is accessible to MIMIT Directorates, and it is used routinely by MIMIT to support policy activity and to prepare reports. For these purposes, the RNA may be integrated with other data sources available to specific DGs. Importantly, since the database contains the name and fiscal code of the firms it can be matched with additional firm-level information enhancing its potential for monitoring and policy evaluations.

The RNA is [available online](#) through an open access platform.⁷ It is, thus, accessible to external stakeholders such as researchers and other public administrations or institutions that may use it for policy analysis and research. Although the database is relatively new, some researchers are starting to exploit it for economic analysis.⁸

In general, the information collected in the RNA offers many possibilities for carrying out policy evaluation and research. For the purpose of this project, the OECD has conducted an initial analysis showing the potential offered by the dataset for monitoring purposes. This analysis shows that important differences emerge between firms asking for one or few state aids and “heavy users” of state aids that exploit several benefits contemporaneously (see Box 2.1), and it provides some insights that could be used by policy makers to enrich the policy design process.

The ability of the RNA to monitor all State aids can also enhance policy evaluation studies. Assessing the impact of a policy by considering its complementarity with other policy schemes can, for example, show whether impacts are more relevant when firms combine multiple measures. Moreover, by excluding firms that benefitted from other aids when building control groups (i.e., firms that do not benefitted from the policy but have similar characteristics to treated firms) it is possible to control for confounding factors that may affect the outcomes of interest and, thus, improve counterfactual analysis. Finally, the RNA allows deriving granular information on the amount of grants received by firms.

7. The database is available at https://www.rna.gov.it/sites/PortaleRNA/it_IT/open_data, where monthly datasets can be downloaded in .xml format.

8. For example, recent analyses have used the RNA to characterise the beneficiaries of State aids over the period 2018-19, showing that industrial policy in Italy is highly fragmented, and that beneficiaries are mostly SMEs and firms located in the South (Albanese et al., 2021^[83]).

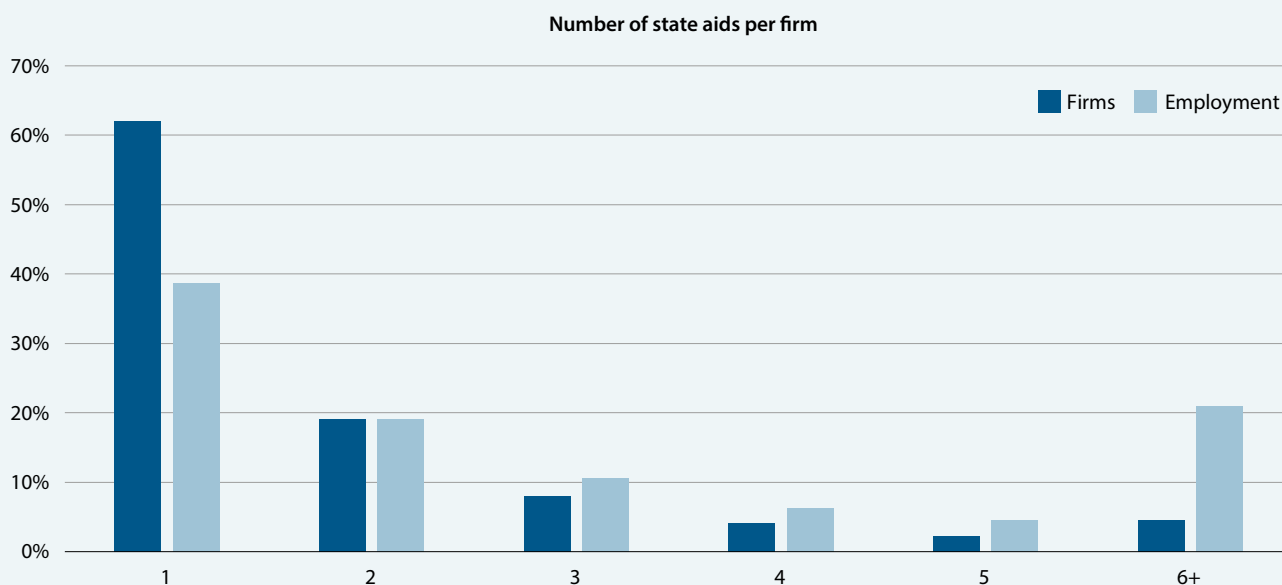
BOX 2.1: CHARACTERISING FIRMS' USE OF STATE AIDS USING THE RNA DATABASE AND FIRM BALANCE-SHEETS INFORMATION

The RNA database allows to have a broad picture of Italian State aids, with respect to allocated resources, managing bodies, and normative bases. If matched with firm-level characteristics, it allows to assess the characteristics of beneficiaries. One key feature of the RNA database is precisely its ability to track firms that accessed multiple policy measures. By exploiting this information, the analysis of the RNA can also allow to characterise the use of policy mixes by firms.

To provide some initial evidence on this regard, the OECD has conducted a preliminary analysis, matching the RNA database with firm balance-sheet information from Orbis for the year 2018.¹ Figure 2.3 shows that most firms registered in the RNA receive only one State aid (62%). These firms are generally smaller than those receiving more than one aid, employing less than 40% of workers in beneficiary firms. The determinants of these differences can be analysed by estimating a linear probability model of a dummy equal to 1 if the firm has been granted more than one State aid (= 0 if the firm has received only one State aid) on a set of firm observable characteristics (size, age, labour productivity) and sector-region fixed effects. The analysis is restricted to limited liability companies, because it is not possible to compute labour productivity for other type of firms. Limited liability companies in 2018 were around one-third of all active firms in the Italian business sector, and they generated over 70% of its aggregate value added.

Additional analyses that the Centro Studi could conduct may allow to further characterise how firms use different combinations of policies, to highlight patterns of complementarity or substitutability across policies, and provide useful information to policy design and evaluation.

FIGURE 2.3. Distribution of firms and employment by State aids – 2018



Note: Dark blue bars represent the distribution of firms registered in the RNA in 2018 with positive employment (76% of the total firms in the RNA) by number of State aids granted. Light blue bars represent the share of employment in 2018 by number of State aids.

Source: OECD elaborations on RNA (extracted January 2022) and Orbis data.

¹ The analysis focuses on firms with positive employment (76% of total number of firms in RNA).

The RNA represents a clear improvement in data integration with respect to what was previously available: the opportunity to have information on the joint use of different State aids by the firm – for instance – may allow to improve our knowledge of complementarities and substitutability between policies. However, the database has still important limitations that affect its potential use for policy analysis and impact assessments:

- **Scope.** Since the database strictly focuses on State aids defined according to EU legislation,⁹ it excludes other relevant incentives administered by the Ministry, such as those financed through general taxation (e.g., the tax credits for R&D investments; investments in fixed assets, “ex-superammortamento”); and investments in Industry 4.0 assets, “ex-iperammortamento”). Nonetheless, these incentives are key elements of the Ministry’s industrial policy and are also extremely likely to be used by firms as complementary or substitutes to other State aids.
- **Coverage.** For State aids monitored by the RNA, information remains extremely limited and coarse. Many State aids provide different policy options in relation to different activities or investments performed by the firm. For example, the “Nuova Sabatini” aid – that support firms’ investments – provides for two different types of subsidies, higher for “4.0” investments, lower for “ordinary” investments. The RNA, however, does not provide details on these subcategories, while generally providing the total amount of subsidy granted to firms for each policy measure.
- **Usability.** The database does not collect any information on firms that applied for policy measures receiving a denial or a revocation, nor does it collect information on the ranking of applicants, if the Law provides it. This information would be particularly useful to perform policy evaluations, as it would help in identifying credible control groups for counterfactual analyses. Moreover, additional information would also help in understanding the effectiveness of policy delivery, e.g., in terms of understanding whether the intervention was delivered as intended and to what extent it reached all the firms that it was intended to.
- **Visualisation.** Despite being publicly available, the online database presents some limitations in terms of data systematisation and visualisation. Indeed, monthly updates consist in additional, separate, file uploads, whose structure (e.g., available variables) may differ over months and years. Consequently, aggregating the information included in the database over the available period may result complex.

Because of these aspects, the scope and depth of analyses that can be conducted using only the RNA (possibly coupled with balance-sheet information on firm performance) is limited.

Nuova Banca dati Agevolazioni (NBDA)

The Ministry has recently created a new database (“Nuova Banca Dati Agevolazioni” or NBDA) that integrates the information collected in the RNA with data on aid payments (only for the State aids financed by the Ministry). The database was created in 2021 to monitor the efficiency of the administration process and to visualise aggregate information by policy measures (including, e.g., the total amount of resources spent, respective number of beneficiaries, etc.). The NBDA includes data on MIMIT State aids’ payments since 2017.

The NBDA represents an important step forward in the data collected by the Ministry. The possibility of monitoring both the granted amount and the actual disbursements (with their date) has indeed manifold advantages. For example, the amount of aid initially granted may differ from the amount received (or used) ex-post by the firm. This information might shed light on a potential not efficient use of public resources. On the other hand, this information might allow to better characterise firms’ behaviour, by examining, e.g., why companies might use only a part of the total resources granted and for what reasons.

Nonetheless, the NBDA presents some limitations in the data collection process, resulting in the NBDA being currently incomplete and not up to date. These limitations are due to a specific feature of the database updating process: while information on the time of granting must be registered – in the case of RNA – for the aids to be effective and paid, there is no legal enforcement or automatic mechanism to register information on payments in the NBDA. Thus, the data updating process may be delayed. Moreover, differently from the RNA, the NBDA is currently not publicly available: the database remains accessible only to some designated personnel within the Cabinet, the General Secretariat and the various DGs. Finally, since the NBDA is directly matched to the RNA, it shares the same limitations in terms of coverage and lack of information.

9. Art. 107(1) TFEU.

Additional datasets available within the Ministry to monitor aids and incentives

In addition to the RNA and NBDA, the Ministry collects **specific detailed datasets** on the beneficiary firms and the characteristics of the incentives for all the **measures it manages directly**.¹⁰ By integrating the information already accessible in the RNA and NBDA, these datasets allow to conduct more detailed and heterogeneous analyses.

A complete list of all additional data available across DGs in the MIMIT is currently not available and Chapter 4 discusses how collecting this information is a key step in developing a structured and comprehensive data-lake to support policy making.

Examples of key databases that have been identified include:

- Specific datasets on State aids collected by DGIAI. For example, the Division VI is responsible for the “Nuova Sabatini” and collects detailed data on its beneficiaries.
- **Micro-data on beneficiaries of incentives not classifiable as State aids** – which is not monitored in the RNA and NBDA. One example is the list of firms that benefitted from the Temporary Export Manager (TEM) voucher, a subsidy to acquire consultant services to support the firm’s trade activity.¹¹
- **Data on potential beneficiaries of incentives for a specific group of enterprises**. For example, DGPIIPMI collects and regularly updates the list of Innovative SMEs and start-ups that are registered in the Special section of the Italian Business Register (*Sezione Speciale del Registro delle Imprese*). The registration is a pre-requisite to access the Ministry’s policy package designed for these types of firms.
- **Data on the beneficiaries of the National Operational Programmes (PON) “Enterprises and Competitiveness” and “SMEs Initiative” 2014-20**, financed with the European funds. Until mid-2022, the DGIAI Division IV provided this data on the Open Data portal PON Imprese e Competitivita’.¹² Following the EU Reg. n. 1303/2013, the data were previously available in CSV format (as well as via web API). As of 2023, the website has been down and inactive for several months.

The use of timely surveys to monitor beneficiaries of MIMIT incentives

Survey data are crucial to monitor Ministry’s policies and improve ex-post evaluation. Indeed, as discussed in Chapter 4, they may provide timely information with a shorter delay with respect to administrative data, and they may complement administrative data with “softer” and more detailed information on firm’s activity. However, these potential benefits may be exploited only if the survey is carefully designed and implemented. Over the past years, the Ministry has collected survey data to monitor several of its policies. The implementation of the survey is usually entrusted to an external agency and/or conducted with the advice of external institutions.

The OECD does not have information on all the surveys administered by the Ministry over the past few years. However, the OECD is currently aware of three surveys recently conducted by the MIMIT:

1. the **Survey on Innovative Start-ups**, which was administered by MIMIT in 2016. It focused only on beneficiary firms (with a response rate of around 40%) and was implemented by an external agency. The survey collected information on pre-policy firm composition as well as the use of policy tools included in the Start-up Act. Data were later exploited in the evaluation of the Start-up Act conducted by the OECD and published as OECD Policy Paper in 2018 (Menon et al., 2018_[3]).

10. In general, detailed information on incentives – and additional with respect to those collected within the RNA and NBDA – are collected by managing entities of the measure (“soggetto gestore”). The managing entity can be either the Ministry or an external entity. Data collected by external entities will be discussed in the next section.

11. It is worth noting, however, that the Ministry was the managing entity for the first two TEM voucher calls (2015-2017), while now (and in the last call, in 2021) the voucher is under the Ministry of Foreign Affairs and the granting entity is Invitalia.

12. <https://www.ponic.gov.it/opendata/>

2. the **Survey to monitor the tax credit incentive for equity investment in innovative SMEs**. This conducted in 2019 by the DGPIIPMI (Division VI) with the support of economists at the Bank of Italy, and it was implemented by an external agency. It was designed to support the in-itinere and ex-post evaluation of the policy. The survey collects both quantitative (e.g., turnover, R&D expenditures) and qualitative indicators (e.g., innovative activities, expectations, propensity to raise equity or debt), including whether the firm is aware of the incentives and would like to benefit from the policy. The survey collected information on a sample of more than five thousand firms, both beneficiaries and not beneficiaries of the policy (identified with similar characteristics).
3. the **Survey “Voucher for Internationalisation”**. This survey was conducted in 2017 by the former DG for Internationalisation Policies (later moved to the Ministry of Foreign Affairs). The survey, administered by an external agency, sampled beneficiary firms (with a response rate of around 80%), and focuses on firm’s ex-ante expectations regarding the policy.

In two out of three of the cases considered (surveys 1 and 3), only beneficiary firms are sampled. Moreover, in these cases, the survey provides very detailed information about firm use of the policy tools, but do not collect information on firms’ characteristics and activities that may be useful to better perform an impact evaluation. Finally, in all three cases, no follow-up surveys have been administered, so that the information remains tightly linked to one specific wave of the policy, with limited capability to contribute to improving the policy design.

More aggregate data available within the Ministry

In addition to microdata on incentives and State aids, the Ministry collects more-aggregate data that might complement the analysis and improve policy evaluation (including e.g., data on innovative activities and sectoral-local information). A complete list of all these data is currently unavailable. However, during the fact-finding missions the Ministry has kindly provided information on some of these datasets.

For example, the DG Intellectual Property (DGUIBM) has aggregate data (at province and sector level) on counterfeiting products, including information on the number of pieces seized, type of product, number of seizures and total value of seizures (Banca Dati Iperico). This data is obtained from the Italian Finance Police (Guardia di Finanza) and the Italian Custom Agency, and it has been exploited by the OECD Illicit Trade paper (2018_[4]). Moreover, the DG collects data on patent applications (patents, trademarks and design applications at the Italian Patent and Trademark Office (UIBM)), including biotechnological inventions and patents from universities. Meanwhile, the DG Internal Market and Competition (DG MCTCNT) collects statistics on the characteristics of wholesale and retail trade sector (Osservatorio Nazionale Commercio), and real-time data on fuel prices and stations (Osservatorio Carburanti).

The common strength of the datasets is their public accessibility. Yet, it remains unclear how these data can be integrated with other available data sources of the Ministry (or external datasets), and whether e.g., it is possible to link them at the firm-level.

Existing relevant datasets outside MIMIT on the beneficiaries of the incentives and firms’ characteristics

Data on MIMIT incentives managed or collected by other entities

In addition to the data available within the Ministry, other entities that manage State aids collect firm-level data on beneficiaries. Among the entities managing MIMIT incentives, the main agencies in charge of MIMIT policies are Invitalia S.p.A – the National Agency for Inward Investment and Economic Development – and Medio Credito Centrale S.p.A. State aids are also administered by other Ministries (e.g., the Ministry of Economics and Finance, the Ministry of Infrastructures, and the Ministry of Culture) and Regional agencies.

In particular, Invitalia holds a large amount of detailed microdata on MIMIT incentives, but most of them are not currently shared with the Ministry. One recent investment in data collection that Invitalia is undertaking – under the broad project called “Minerva” – is the restructuring of its Data Lake with the aim of enriching the integration of data on all incentives managed by the Agency. The Data Lake will contain a rich set of granular information on the entire history of beneficiary firms (including, e.g., if firm has invested training activities etc.), and will be made accessible to the Ministry in 2023.

Among other entities that managed MIMIT incentives, **Mediocredito Centrale S.p.A** is responsible of managing the Public Guarantee Scheme (“Fondo di Garanzia”), and thus collects detailed data on its beneficiaries. These data are also available online¹³ and are up to date.

For what concern the MIMIT tax credit incentives, the data on beneficiary firms, the amount invested, the deduction and additional information are collected by the **Italian Tax Agency**. The Italian Tax Agency transmits data about fiscal State aid to the RNA two years after the enterprises benefit of them, because of the peculiar functioning of fiscal aids.

Managing entities already transmit a limited set of information to the RNA and NBDA (only for State aid policies). However, more-detailed data on beneficiaries of MIMIT policies are not currently shared with the Ministry, and their integration could provide important insights and additional information to conduct policy analyses and evaluations. Strengthening data sharing agreements with these entities is thus crucial to conduct in-depth policy analysis and evaluations.

Data on firm characteristics and performance

To perform a sound policy evaluation and support evidence-based policy making, it is key to complement data on State aids with detailed and comprehensive information on firm’s characteristics and activities. For this purpose, the data capability of the main Italian statistical institutions is extremely developed, and the various databases are generally easy to integrate. This section summarises the main characteristics of the more relevant databases available.

The backbone of the sources of firm-level data is the **business registry**, which in Italy is collected by the Chambers of Commerce and administered by InfoCamere (a public firm acting under private law). This database collects detailed information on each establishment active in the business sector, together with information on ownership and control of each Italian private firm.

Balance-sheets information are also collected by the Chambers of Commerce, but they are by and large limited to limited liability companies which represent less than 40% of total firms, while accounting for over 70% of total value added of the business sector. ISTAT complements this information with data from the Tax Agency (and other sources) to compute some key **income statement variables** (such as turnover and value added) for firms that do not file balance-sheets.

Employment information mainly come from the Italian Social Security Institute (INPS). INPS collects detailed matched employer-employee data on labour input and wages, as well as information on social security contributions, unemployment benefits, sick leaves, and use of active labour market policies. INPS data are integrated into the data infrastructure of both InfoCamere and ISTAT.

The integrated database administered by ISTAT that **combines both income statements and employment data** is called ASIA-FRAME. The ASIA-FRAME database has been carefully crafted to allow studying micro-to-macro linkages of key economic variables. It also provides aggregate statistics that are internationally comparable, as the definition of firm, its birth and death follows the standards set by Eurostat. Further developments of the ASIA-FRAME database include the matched employer-employee data (FRAME-Occup) and the match between firm-level data and product-level data on imports and exports from the Custom Agency (ASIA-COE).

Besides administrative data, **large surveys** are also key to analyse more in-depth firm activity. ISTAT conducts a set of surveys to analyse innovative activity and technology adoption. For example, the periodic Multi-purpose Survey, administered by ISTAT, and the yearly Census surveys, provide extremely rich information on firm’s practices, knowledge, organisational structure, and strategies for a large and representative sample of firms. All ISTAT surveys can be matched with the ASIA-FRAME database.

13. <https://www.fondidigaranzia.it/amministrazione-trasparente/>

Another important firm-level survey is the yearly Survey on Industrial and Services Firms (INVIND), administered by the Bank of Italy. INVIND provides important and timely information (the questionnaire is administered at the end of Q1 in each year) on a representative sample of around 5 thousand firms. Information elicited include their current and expected main balance sheets variables and their expectations about future market conditions. The Bank of Italy also administers the national **Credit Registry**, which contributes to the European-level Registry ANACREDIT. ANACREDIT provides detailed information on each credit contract stipulated between banks and firms, including the amount granted and disbursed, its price, maturity, and collateral value.

Recently, the Tax Authority has started collecting **VAT data on firm-to-firm and firm-to-consumer transactions**. While its use is still extremely limited, carefully exploiting information on firm-to-firm transactions, paired with balance-sheet and other information may allow to considerably improve the design and targeting of industrial policies.

OVERVIEW OF PAST AND ONGOING MONITORING AND EVALUATION EXERCISES OF MIMIT POLICIES

This section presents an overview of the activities aimed at quantifying the impacts of the policies under the responsibility of the Italian Ministry of Economic Development (*Ministero per lo Sviluppo Economico*, MIMIT). First, it reviews the framework for regulatory impact assessments at MIMIT. Then, it looks at the monitoring of some of the policies under MIMIT responsibility. Finally, it overviews the evaluations conducted by MIMIT, as well as those conducted outside MIMIT that the Ministry has used to inform policy improvements. In particular, the section reviews the use of empirical analysis to quantify the impacts of the Public Guarantee Scheme for Small and Medium Sized Enterprises (*Fondo di Garanzia*) and the so called “Start-up Act”. These examples provide useful insights on the current and potential capacities and challenges for quantifying policy impacts at MIMIT.

MIMIT and the framework for regulatory impact assessments: AIRs, VIRs and current gaps

Legal frameworks are important enablers of the emergence of a culture of evaluation in the public administration (Box 2.2). In Italy, all legislative acts (with some important exclusions) require a regulatory impact assessment as expressed by law n.246 of 28 November 2005 and the decree law n.169 of 15 September 2017. Regulatory Impact Assessment is a systemic approach to critically assessing the positive and negative effects of proposed and existing regulations and non-regulatory alternatives (OECD, n.d.^[5]).

While MIMIT does perform some form of Regulatory Impact Assessments (in Italian AIR), these suffer some substantial gaps, which is not an unusual situation across the Italian administration. MIMIT performed 15 AIRs in 2020, against 4 in 2019 (DAGL, 2020^[6]). Inside MIMIT, AIRs are produced directly by D.G.s, and are not conducted in a systematic coordinated way given the decentralised nature of the system until now. MIMIT’s Legislative Office is responsible, *inter alia*, for collecting the AIRs and assessing that the latter comply with the related legal standards.

According to Italian law, including the law 2005-246, completed by the Decree Law 2017-09-15, ex-post assessments (*valutazione d’impatto della regolamentazione*, VIR) have to be performed after two years from the implementation of a legislation. The guides and directives prepared by the Presidency of the Council offer guidance for implementation. The phases and fundamental steps to perform these analyses are established by Art.12 and Art.13 of the Prime Ministerial Decree n.169, 15 September 2017. Administrations are required to present a biannual programme of ex-post evaluation to the Prime Minister Legislative Office (DAGL) in which the evaluations that will be performed are listed. However, five years after the decree MIMIT has not performed any VIR yet.

Further to the first introduction of these elements in the Law n. 50 of 8 March 1999, these instruments are often used as a purely legalistic exercise rather than taking advantage of them to evaluate economic and regulatory impacts. AIRs are more common than VIRs as they are required to present a legislation to the Council of the Ministers while the production of VIRs is not subject to enforcement. As reported by the Legislative Affairs Department of the Presidency of the Council (DAGL), 48% of AIRs are considered inadequate when sent to the DAGL for the first review. The most common problems that characterise AIRs are the inability to assess the impact (more than 60% of cases) and the lack of consultations (more than 40% of cases). Despite these gaps, some improvements were observed in recent years (DAGL, 2020^[6]).

BOX 2.2: SPAIN: EMBEDDING EVALUATION AND MONITORING IN LEGISLATION

A pioneering example for the advancement of public policy evaluation was recently implemented in Spain, where, in 2022, the Council of Ministers approved legislation aimed at strengthening, systematising and providing stability and quality to the evaluation of national public policies (Gobierno de España, 2022_[7]). Through the law, Spain aims to achieve four objectives: (i) strengthening the public sector; (ii) promoting an evaluation culture; (iii) involving the public sector in its delivery of public service and public value; and (iv) measuring policy impact in a transversal, integral and participative manner, beyond fiscal accountability. To this end, the law lays out a series of principles that require evaluation to be comprehensive, cross-cutting, independent and participatory (Ministerio de Hacienda y Función Pública, 2022_[8]). The text also introduces a systematic evaluation planning mechanism for the entire administration, through quadrennial plans for strategic evaluations that are deemed most socially or economically relevant, and biannual evaluation plans for each government departments (Cribillers, 2022_[9]).

In addition, the law also mandated the creation of a new public body, the Agency for the Evaluation of State Public Policies, which is expected to be in place by the third quarter of 2023. The text also envisions an inter-ministerial collegiate body for public cooperation and participation (*Comisión Superior de Evaluación*) and a General Evaluation Council, which would allow participation of civil society and other interested parties in the design and implementation of evaluations (Gobierno de España, 2022_[7]).

Monitoring

Monitoring aims at measuring the implementation of a policy measure to ensure that planned activities are on track, outputs are delivered and any changes of the underlying conditions are taken into consideration. Monitoring is descriptive and generally relies on administrative data related to the management of the policy and the characteristics of the beneficiaries. It is closely linked to evaluation, as monitoring data are necessary to evaluate the impact of policies. These indicators can also feed into the AIR and VIR processes described above.

MIMIT produces regular monitoring reports on the policies under its responsibility. Some monitoring reports are mandated by law and include descriptive statistics on take-up and beneficiaries. They include:

- **Relazione sugli interventi di sostegno alle attività economiche e produttive:** this is an annual report presenting a stocktaking and monitoring of the business incentives that the national and regional administrations have allocated to productive activities. It presents a detailed analysis of business incentives at the national level (number of applications, incentives granted, incentives disbursed) in the form of ex-post descriptive statistics along different dimensions of the national incentives system. The data to monitor these variables is collected through a dedicated system (*Piattaforma 266*) and is complemented by data from the RNA. The report is prepared by the Directorate-General for Business Incentives (DGIAI) and sent to Parliament by 30 September of every year (MIMIT, 2021_[10]).¹⁴
- **Relazione annuale sullo stato di attuazione della normativa a sostegno delle startup e delle PMI innovative:** this is an annual report monitoring the measures supporting start-ups and innovative firms under the 2012 so called “Start-up Act”¹⁵ (on which see more below). It provides descriptive statistics on the take-up of the measures, including geographical distribution, type of activities, access to credit and capital as well as age and gender of the beneficiaries. It also provides an overview of the main trends in terms of increase of innovative start-ups, employment and investment. The 2021 report presents the results of a qualitative survey conducted in July 2020 by MIMIT on a representative sample of 1,600 cooperatives, to understand the strategies adopted by innovative firms to respond to the COVID-19 crisis. The report also includes the results of a survey conducted by MIMIT in 2019 to monitor the tax credit aimed at supporting investment in innovative SMEs (MIMIT, 2021_[11]). The report is prepared by the Directorate-General for Intellectual Property (DGUIBM) and sent to Parliament by 1 September of every year since 2014 as mandated by the “Start-up Act.” Quarterly monitoring reports on the measures under the “Start-up Act” (*Fondo di Garanzia per le PMI, Startup Visa&Hub*) and data on the features of innovative firms (number and size,

14. Art. 1 of law 266/97, and art. 14 of Law 115/2015 (European Law 2014).

15. Law Decree 18 October 2012, n. 179, Section IX (measures related to the creation and development of innovative start-ups), <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legge:2012-10-18;179vig>

economic sectors and typologies, geographical distribution, employment, turnover, profitability) are produced and available on MIMIT website (MIMIT, 2022_[12]).

Other monitoring activities are conducted internally, as is the case for the data collection activities described in the section of available MIMIT datasets above (Chapter 2).

MIMIT has also acquired additional monitoring responsibilities under the National Recovery and Resilience Plan (NRRP). In particular, MIMIT is responsible for ten investment projects and a reform of the legal framework on industrial property under the NRRP. The overall implementation of these measures will be coordinated with the support of the monitoring office (*Ufficio di Monitoraggio*), which will be in charge of monitoring the NRRP.¹⁶ The Ministry will be in charge of:

- collecting data and periodically monitoring the predisposition and implementation of NRRP projects under its responsibilities at the sectorial and territorial level,
- assessing compliance with planned milestones and targets.¹⁷

Monitoring of the NRRP is accessible on a dedicated website of the Presidency of the Council of Ministers (*Presidenza del Consiglio dei Ministri*, PCM) in the form of a matrix with succinct information on the NRRP objectives, the public organisation in charge of implementation and evidence on the milestones achieved (PCM, n.a.[13]).

Evaluation

Evaluation aims at assessing the actual or potential impact of a policy measure, looking at, among others, whether the objectives of the policy have been effectively and efficiently met, the extent to which impacts are directly or indirectly attributed to the policy, to what extent the policy has created unintended positive or negative side-effects, what policy features have been particularly effective (and what has worked less well). Evaluation builds on monitoring data that should be integrated with other socio-economic data.

MIMIT has been a source of data for a number of evaluations, some of which have used counterfactual methods to quantify policy impacts. These evaluations have been mostly carried out at the initiative of public organisations and academics outside the Ministry, which has used the findings of some of these evaluations for policy improvements. MIMIT commissioned an evaluation based on qualitative and quantitative methods (including counterfactual) of an important business incentive (Nuova Sabatini), which was conducted by researchers outside the Ministry. It has evaluated EU-funded programmes, including through quantitative methods, through Invitalia, an agency of MIMIT, and MIMIT Directorate Generals.

Evaluations have been conducted in the framework of EU-funded programmes, mostly by external researchers

MIMIT's DGIAI has been the implementing authority for the Operational Programme "Enterprises and Competitiveness" 2014-2020 (*Programma Operativo Nazionale Imprese e Competitività 2014-2020 – PONIC*). The programme aimed at strengthening the competitiveness of Italian firms in the global production context (in particular SMEs in Southern Italy) by combining business incentives, support for digitalisation and green transition (Invitalia, 2022_[14]).

As part of the programme, an ex-ante evaluation was conducted by a group of external consultants. The evaluation was set up in line with the EC Regulation 1303/2014 and EC guidance on ex-ante monitoring and evaluation (EC, 2014_[15]), and guided the initial implementation phase, with a view of correcting any aspects which could invalidate the subsequent success of the initiative. The evaluation built on a combination of systematic reviews, desk analyses and interactions with key stakeholders, and was mostly qualitative, not building on counterfactual impact evaluations or experimental approaches. The evaluation focused on: i) complementarities and consistencies of the Programme with other relevant instruments at EU, national and regional level; ii) relevance and clarity of the proposed output and

16. Established by MIMIT/MEF Decree, Strutture di missione PNRR MIMIT, 9 September 2021, https://www.MIMIT.gov.it/images/stories/trasparenza/2022/struttura_di_missione_pnrr_decreto_bollinato_signed.pdf

17. Art.8, c.5bis, DL Governance of the NRRP, n. 77/2021, <https://www.gazzettaufficiale.it/eli/id/2021/05/31/21G00087/sg>

result indicators, the statistical robustness, the feasibility of the targets set for 2023 and the relevant milestones; iii) alignment of financial allocations with programme objectives; and iv) programme's governance, monitoring and data collection mechanisms.

In order to assess the effectiveness and efficiency of the PON projects, MIMIT also adopted a PONIC Evaluation Plan (approved in August 2016 and regularly updated) as the reference framework for the evaluation activities, providing for cognitive surveys and analyses, implementation assessments and evaluation of the outcomes. In line with the plan, a few evaluations have been completed. These evaluations were conducted by external researchers. For example, an Italian academic institution, the *Scuola Superiore Sant'Anna*, collaborated with a consultancy (Vision & Value) to conduct a counterfactual evaluation of the effectiveness of the interventions and the efficiency of the implementation tools of selected PONIC incentives.

Counterfactual evaluations conducted outside MIMIT have helped inform the reform of the Italian public guarantee scheme for SMEs

MIMIT manages the *Fondo di Garanzia* (FG), Italy's public guarantee scheme to support SMEs' access to bank credit. The FG started operating in 2000 and the volume of guaranteed loans has grown significantly, especially after the global financial crisis, becoming an important instrument to support Italy's SMEs. New loans guaranteed from 2009 to 2016 amounted to EUR 90 billion compared to EU 11 billion in the period 2000-2008 (Ciani, Gallo and Rotondi, 2020^[16]). In its original design, the FG guaranteed up to 80% of the value of a bank loan (with the value of the guarantee limited to EUR 1.5 million). Banks have to verify the eligibility of the firm through a scoring system that classifies firms as Type 0 (not eligible), Type 1 and Type 2 (both eligible but with different degree of assessment given the different credit scores). Indicators to determine the scoring included ratio of equity and long-term debt on fixed assets; leverage ratio; interest expenses over sales; cash flow over sales (de Blasio et al., 2018^[17]).

Given the relevance of the instrument for the Italian economy and the significant volume of public resources involved, assessing the extent to which the FG targets the "right" firms is particularly important for the policy maker to i) ensure targeting firms that effectively need credit and would not otherwise access credit, and ii) ensure targeting firms that are likely to repay the loan (Manaresi, 2021^[18]).

The evaluation was conducted in 2018 relying on data shared by MIMIT with a team of researchers of the Bank of Italy and the Catholic University of Milan, which complemented the administrative data on the FG with data on firms' balance-sheet acquired through CERVED, an Italian company specialised in collecting and elaborating firm-level data (de Blasio et al., 2018^[17]). The evaluation relies on a regression discontinuity design, looking at firms just below the eligibility threshold and comparing them with firms just above the eligibility threshold. This first evaluation triggered a second evaluation using a similar design applied to data from Unicredit, a large commercial bank (Ciani, Gallo and Rotondi, 2020^[16]). The evaluation found that:

- The FG has been effective in enhancing credit flows; however, the scheme increases the probability that a firm is unable to pay back;
- The impact of the credit guarantee on investment and sales has been limited (and mostly used to finance inventories and trade credit);
- Banks tend to grant the credit guarantee to less risky firms with a long-standing relationship with the bank.

In 2019, MIMIT used the results of the evaluations to significantly change the FG's scoring system:

- The rating has been reviewed to calculate the probability of default of the applicant. Ratings range from 1 to 12 and are grouped in five classes of creditworthiness: 1. Safe; 2. Solvent; 3. Vulnerable; 4. risky; 5. not creditworthy. Firms falling in the fifth class are not eligible.
- The guarantee coverage has been lowered for safer firms and short-term loans.

- The rating model is not applied to start-ups for which the evaluation is based on the business plan and the coverage ratio remains at 80% (Ciani, Gallo and Rotondi, 2020_[16]).

A potential limitation of the new system is that the default risk is also used to assess the difficulties of the firm to access credit, with the risk of continuing to exclude some firms that would need to access credit (Manaresi, 2021_[18]). This points to the need to plan already further evaluations once there is sufficient evidence.

The monitoring and evaluation provisions inscribed in the “Start-up Act” are facilitating the evaluation and improvements of the policy

The “Start-up Act” is the policy framework implemented by MIMIT to support innovative firms across all sectors through support until the fifth year since incorporation. It includes a bundle of provisions that are expected to support innovation, including policies for facilitating entry and exit to the market, tax incentives, labour and remuneration provisions, incentives for equity crowdfunding.

The enabling legislation passed in 2012 includes specific provisions aimed at facilitating the monitoring and evaluation of the economic impact of the policy measures, taking into consideration data requirements, resources and access, as well as obligations to feed monitoring and evaluation results in the policy-making process. Specifically, the provisions include:¹⁸

- A permanent monitoring and evaluation system set up at MIMIT to assess the impact of the measures on growth, employment and innovation, building also on data from the National Statistical Institute (*Istituto Nazionale di Statistica*, ISTAT) and other public institutions through the *Sistema Statistico Nazionale* (SISTAN).¹⁹
- Monitoring reports on the state of implementation of the measures mandated at least annually to identify potential corrections and adjustments.
- Databases to be developed by ISTAT and accessible free of charge to inform the monitoring and evaluation of the measures. ISTAT is allocated resources (EUR 150 000 in 2013, 2014 and 2015) to set up the databases.
- An annual report to be presented by MIMIT to Parliament on the implementation and impact of the measures to inform any necessary modification of the law.

The policy also includes features that facilitate data collection, in particular:

- A special section dedicated to start-ups in the *Registro delle imprese* (Business Registry) managed by the Chambers of Commerce with detailed information on beneficiary firms (including date of incorporation, activity code and description of activities, legal status of the firms, management). Data are updated weekly and freely accessible (<https://startup.registroimprese.it/isin/home>).
- In line with the monitoring requirements, MIMIT and ISTAT conducted a survey of the start-ups accessing the measures under the policy between March and May 2016. The survey covers 44% of registered start-ups (as of December 2015) and addresses questions related to the background, education and family situation of the start-ups’ founders, shareholder structure, type of innovative activities and knowledge sources, which specific policy measures the firms have used (MIMIT and ISTAT, 2018_[19]).
- The Business Registry and survey dataset are linkable through a unique firm identifier which also allows the dataset to be linked to other administrative and firm-level data sources.

18. Law Decree 18 October 2012, n. 179, Section IX (measures related to the creation and development of innovative start-ups), art. 32 (communication and evaluation of the impact of the measures), <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto:legge:2012-10-18;179/vig>

19. SISTAN is the Italian network of public and private entities which provides official statistical information to national entities and international organisations. Established in 1989, it is managed by ISTAT.

The provisions in the law and the production and availability of data have facilitated the evaluation of the “Start-up Act”, either independently from MIMIT or with inputs from MIMIT (but without a direct role in conducting the evaluation).

A team of the Bank of Italy conducted a first evaluation already in 2016, using data from the Business Registry and firms’ balance-sheet data from CERVED. Using a propensity score matching control sample created before the enactment of the “Start-up Act”, the evaluation found that the incentives to increase bank loans and those for equity investors had been efficient in increasing external funding for innovative start-ups. The study also found that strict criteria for innovation were crucial to target very innovative start-ups (Finaldi Russo, Magri and Rampazzi, 2016^[20]). However, this finding was not taken into consideration when the benefits for innovative start-ups were extended to innovative SMEs in 2015.

Another evaluation was conducted in 2018 by the OECD, the Bank of Italy and the *Scuola Superiore Sant’Anna* with inputs from MIMIT. The evaluation focuses on start-ups and includes a counterfactual assessment (difference-in-difference with an instrumental variable used to control for reverse causality) of the impact of the policy using Business Registry and survey data linked to balance-sheet data from Orbis, a commercial balance-sheet repository maintained by Bureau Van Dijk. The study finds that the policy is effective in increasing beneficiary firms’ revenues, value added and assets relative to similar start-ups not benefitting from either credit guarantees (see above) or support for equity financing.

Some of these impact evaluation studies are also presented in the Annual Reports to Parliament on innovative start-ups and SMEs.²⁰ For example, the 2021 report mentions the results from a Bank of Italy/OECD propensity score matching evaluation of the impact of MIMIT incentive for equity investment in innovative SMEs, summarising the findings of the evaluation. It also reports other evaluation activities conducted under the PONIC. References to these evaluations could inform suggestions on policy improvements.

MIMIT commissioned a quantitative and qualitative impact evaluation of a long-standing incentive aimed at supporting SME investment

Except for the cases of evaluations conducted in the framework of EU-funded projects, the evaluations described above were conducted outside MIMIT, which had in some cases provided access to data and facilitated the evaluation. In 2017, MIMIT commissioned an evaluation of the so-called “Nuova Sabatini”, an incentive to support SMEs’ investment in tangible and intangible assets that is particularly important in the context of the Italian economy (Box 2.3).

The evaluation was commissioned through Invitalia, which is the agency of the Ministry responsible for managing the implementation of all incentives supporting firms under the responsibility of MIMIT. Invitalia also supports the public administration in implementing national and EU-funded projects. The evaluation sought to address direct, indirect and negative effects of the incentive as well as adequacy and heterogeneity (Table 2.1).

20. <https://www.MIMIT.gov.it/index.php/it/impresa/competitivita-e-nuove-imprese/start-up-innovative/relazione-annuale-e-rapporti-periodici>

BOX 2.3: THE SME INVESTMENT INCENTIVE SCHEME “NUOVA SABATINI”

Created in June 2013, the incentive facilitates the financing of tangible and intangible assets for SMEs. It provides firms with a contribution equal to the interests of a five-year loan of the investment amount at an interest rate of 2.75% (increased at 3,575% for 4.0 investment in 2019). The incentive can add to other incentives, within the EU State Aid rules.

The “Nuova Sabatini” replaced a similar instrument created in 1965 in the face of a 30% fall in gross fixed investment in the period 2007-2015 (Eurostat). The stated objective was to reverse this trend by supporting the replacement of obsolete machinery (tangible assets) and the use of new technologies (intangible assets). Funding for the incentive amounts to EUR 1.3 billion through 2023. It has been estimated that from April 2014 to October 2018, the contributions granted by MIMIT, net of total revocations, withdrawals and cancellations, amounted to EUR 977 million (amounting to 86.6% of the contributions which were originally committed) for a total number of about 52 500 granted requests for support.

Source: MIMIT: Beni strumentali – Nuova Sabatini, <https://www.MIMIT.gov.it/index.php/it/incentivi/agevolazioni-per-gli-investimenti-delle-pmi-in-beni-strumentali-nuova-sabatini>

TABLE 2.1. Research questions for the evaluation of the “Nuova Sabatini”

Which impacts?	What to evaluate?
Direct impacts	Were tangible and intangible assets modernised? Did firms become more efficient? Did firms have better access to credit?
Indirect impacts	What was the impact on firms’ performance? Was there an impact on employment? Has there been an impact on the probability of firm’s survival? Was there an impact on the firms’ capacity to innovate?
Heterogeneity of impacts	Are impacts differentiated by firms’ size and location? Do impacts change if the incentive is cumulated with other incentives? What type of firm has benefitted the most? On which basis have banks selected firms?
Negative impacts	What are the negative effects on non-beneficiary firms? What are the negative effects on beneficiary firms? Why would firms choose the incentive rather than the tax credit for investment in Southern Italy?
Proportionality and adequacy	Have resources met the needs? Was the incentive commensurate with the objectives?

Source: (Invitalia, MET and CSIL, 2020_[21])

The evaluation was conducted by external providers (MET srl and *Centro Studi Industria Leggera*) and included:

- Theory-based evaluation to “reconstruct” the intervention logic and highlight the expected effects of the incentive;
- Counterfactual analysis (using a combination of the difference-in-difference approach and propensity score matching) to assess the actual impact of the incentive;
- Bayesian network analysis to assess the causal channels of the impacts produced by the incentive (Invitalia, MET and CSIL, 2020_[21]).

The evaluation found that the impact of the incentive is largely positive in terms of modernising assets and firms’ efficiency. A smaller effect is noted on employment and access to credit. The impact is greatest among micro-enterprises especially in the South, which tend to be the most fragile (less so for the more dynamic sectors in the most productive regions) (Invitalia, MET and CSIL, 2020_[21]).

Overall, the evaluation makes extensive use of data (including through two qualitative surveys among beneficiary firms) and applies counterfactual methods to quantify impacts. The evaluation also leaves a series of open questions for the policy maker, including the extent to which the incentive can be better targeted to make the best use of resources and how the interactions with other incentives can be better exploited to better target the menu of incentives to firms. These are in part questions that would need to be addressed and explored in greater depth “in-house” to provide the policy maker with options for the improvement of the incentive. For this reason, as part of this OECD-EC project, the Nuova Sabatini was selected to be used as a sample policy to overview the process of performing a pilot evaluation at MIMIT (Box 2.4).

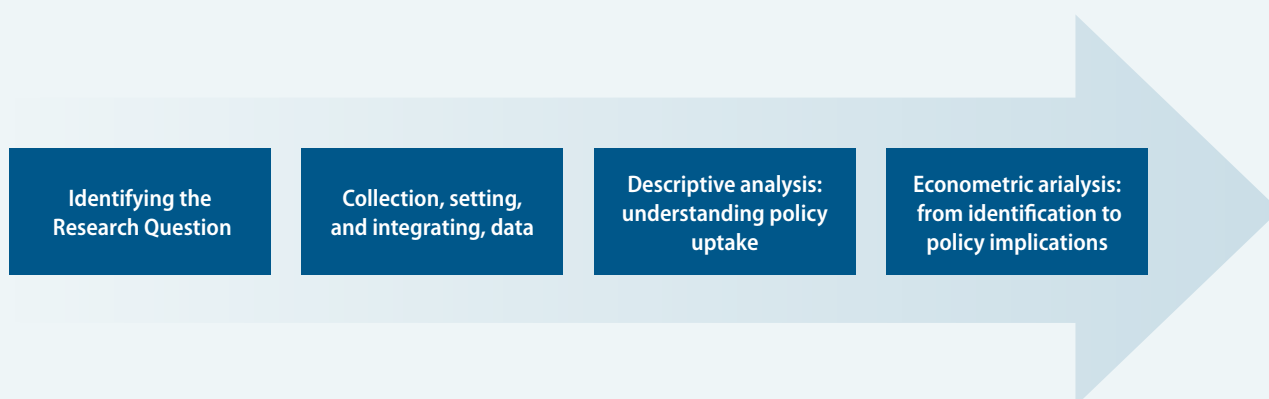
BOX 2.4: CONDUCTING A SAMPLE POLICY EVALUATION AT MIMIT: THE NUOVA SABATINI PILOT EVALUATION

As part of the project, the OECD, in collaboration with the Centro Studi, conducted a sample impact evaluation of the Nuova Sabatini, exemplifying, through three interactive webinars, the different steps needed to perform an impact evaluation in the domain of MIMIT from beginning to end.

Before starting the evaluation, the OECD reviewed available datasets to identify whether the policy could be evaluated and what datasets should be obtained to this end, as well as from which institutions. It was identified that some datasets were already available within MIMIT (e.g., on the take-up of cumulable incentives: Registro Nazionale degli Aiuti), while others could be obtained from external sources (e.g., data on loan interest rates Dataset Taxia from Bank of Italy; data on financial statements of capital companies from Infocamere; data on employment from INPS or ISTAT etc).

The sample ex-post impact evaluation included initial descriptive statistics to understand how many and what kind of firms benefited from the policy, followed by quasi-experimental counterfactual analysis (*Propensity-score matching* and *Difference-in-differences*) to gauge the effectiveness of the policy.

FIGURE 2.4. The process of conducting the Pilot Evaluation



Source: OECD Presentation on Pilot Evaluation.

The process and results of the pilot evaluation were disseminated through three thematic workshops for the MIMIT Centro Studi and interested Directorate Generals, held in December 2022. The exercise drew lessons for MIMIT on how the Centro Studi could conduct similar evaluations, providing guidance on the type of methodologies that would be used and the data that would be needed, with indication of the possible institutions with which it would be most helpful to partner for a full-scale evaluation of the Nuova Sabatini. In particular, the activities were structured as follows:

- **A first workshop** focused on how to collect, harmonise and integrate the different data sources used in the analysis.
- **A second workshop** presented descriptive statistics on the take up of the policy (and its heterogeneity by firm's characteristics), and how firms cumulate it with other incentives
- **A third workshop** evaluated the impact of the policy using matching and diff-diff strategy and quantify the impact.

Moreover, the OECD shared with MIMIT comprehensive materials on the pilot evaluation, including relevant background materials supporting the evaluation as well as the statistical codes to implement the analysis. Results were also disseminated to the European Commission in a dedicated meeting of the Advisory Group of the project in January 2023.

MIMIT is expected to contribute to an in-depth evaluation of a sizeable NRRP intervention aimed at strengthening business digitalisation

Together with the Ministry of Economic and Financial Affairs (MEF) and Bank of Italy, MIMIT is part of a Scientific Committee tasked with evaluating the economic impact, the effectiveness and the efficiency of various NRRP interventions for business digitalisation, including fiscal incentives, implemented in the framework of the NRRP investment “Transition 4.0.” The intervention allocates approximately EUR 18 billion to this end (or 8% of the entire NRRP envelope), articulated in five tax credits (one for tangible investments, two for intangible investments, one for R&D and one for training). The NRRP sets the following targets:

- 69,900 tax credits granted on the basis of 2021-22 tax returns by Q2 2024;
- 111,700 tax credits granted on the basis of 2021-2023 tax returns by Q2 2025 (MIMIT, n.d.^[22]).

The impact evaluation is expected to be carried by end 2025. No activity has been yet carried out by the Scientific Committee.



3. Strengthening the Centro Studi's governance drawing on international best practices

This chapter presents some best practices identified in the area of governance that can help strengthen the Centro Studi's strategic framework to foster evidence-informed policy-making inside MIMIT. It concentrates on four fundamental steps: identifying the stakeholders and priorities, building internal capacities, producing high quality results and communicating them in an effective way. It does so by using the OECD existing knowledge on evidence-informed policy-making based on best practices adopted in similar contexts and examples of evaluation units in Ministries similar to MIMIT.

SETTING A STRATEGIC FRAMEWORK TO ACHIEVE CENTRO STUDI'S OBJECTIVES

Setting a strategic framework can help to better contextualise the experiences of other countries and use them as a "toolbox" for the recommendations of the project (see Section 1 on Main findings and recommendations). To ensure that policy evaluations will become integrated in MIMIT's policy-cycle, the Centro Studi will have to perform multiple activities. Indeed, a well-functioning evaluation system requires three fundamental aspects: institutionalisation, quality and use (OECD, 2020_[23]) and the Centro Studi will have to invest in all these aspects.

To translate these elements into more practical steps the Centro Studi will have to:

- **Define a clear mission and set specific objectives.** Both the mission and the objectives will have to be in line with the Ministry key strategic objectives stated in the *Atti d'indirizzo*, other strategic documents and the RRP. This will require constant discussions and engagement with decision-makers and Ministers to ensure that the Centro Studi focuses on the most relevant evaluations. This represents the first fundamental step to achieve impact.
- **Once the strategic objectives are clear, the Centro Studi can start producing evaluations.** To produce evaluations, good data and the right skills are necessary. To guarantee a sustainable production of evaluations, data and skills will need to be developed inside the unit. At the same time, these skills will have to be mobilised also inside the different D.G.s and in collaboration with external stakeholders.
- **It is important to ensure that evaluations performed by Centro Studi are of good quality.** This is vital to ensure robustness of results and hence trust and reliability both from the Ministry than from the public.
- **Finally, once the evaluations are produced, the results will have to be communicated in an effective way** to make sure that decision-makers are able to learn and use them to design and guide policy implementation. To enhance transparency and promote an evaluation culture, it is also important that results are published and made available to the public.

To further understand what each of these steps entails the report provides examples from other countries and their ways to define the objectives, build capacities, strengthen quality and communicate both internally and externally.

Countries experiences can facilitate to grasp the elements of success which could be translated in the Italian case. At the same time, organisational structures are often context-related and are strongly influenced by the national institutional culture. Moreover, several evaluation systems tend to be framed with a whole-government approach, presenting some higher level of coordination often steered by the Prime Minister's Office or the Ministry of Finance. These bodies are often in charge of promoting an evaluation culture, creating common guidelines across line ministries and having a coordination role. Still, line ministries themselves can have significant autonomy in deciding how to organise and produce evaluations. This is indeed common in countries with more decentralised evaluation systems like Germany and the Netherlands, while countries with a more centralised approach can leave significant space to

line ministries, as in the United Kingdom. Creating a common evaluation system across the different Italian Ministries could help ensure a better systematisation of the evaluation and a share of best practices and knowledge but it is beyond the reach of the current report. In the meantime, there are already specific activities that MIMIT can undertake in order to foster evaluation at the internal level.

IDENTIFYING KEY STAKEHOLDERS AND THEIR PRIORITIES

To be used, evaluations need to capture the interest of policy-makers and of influential external stakeholders. An important way to assure interest in the evaluations is to perform them in the areas that are considered pivotal for the Ministry and for other key stakeholders. This means that evaluations must be utilisation-focused and planned accordingly to the MIMIT strategic objectives. This will of course include also the broader National Resilience and Recovery Plan. OECD countries have adopted a series of mechanisms to ensure that most important evaluations are prioritised. In particular, countries are using stakeholder engagement and evaluation plans as instruments to ensure production of the right evaluations at the right time.

To increase chances of impact, it is important to produce utilisation-focused evaluations, where the needs of the primary users are taken into account (Patton, 1978_[24]). In order to create utilisation-focused evaluative processes, two aspects are particularly important: stakeholder engagement and planning.

Including both internal and external stakeholders is a way to generate broader consensus and increase legitimacy (OECD, 2016_[25]). Moreover, evidence shows that policy-makers are more likely to use evaluation results when they come from trusted sources rather than from formal sources (Oliver et al., 2015_[26]; (Haynes et al., 2012_[27])). In addition to including a variety of stakeholders, it is also important to include them in phases of the evaluations where they can have a real contribution. Often, a best practice is represented by involving them from the early stages of the evaluation. An example of stakeholder engagement comes from the German Federal Ministry of Economic Affairs and Energy (BMWK), which, despite externalising evaluations to third parties, remains active in defining the scope and design of the evaluations to assure its relevance for the Ministry.

Another important way to create utilisation-focused evaluations is to perform these when the users need them. This can only happen with a good planning of resources and requires the creation of an annual or multiannual evaluation plan. These plans are intended to establish timelines for the evaluations to be carried out each year, with indication of databases to be adopted; estimation of expected costs and a concrete workplan, with indication of the composition of responsible evaluating teams and specification of the different responsibilities. This is a widely used practice in OECD countries (See Box 3.1 for example). Decision-makers should participate to the creation of such a document to make it tailored to their needs.

STRENGTHENING CAPACITY FOR ANALYSIS

Policy evaluations require significant skills and data. This demands an investment that can be challenging. Indeed, the lack of technical skills and infrastructures is considered one of the main obstacles of evidence-based policymaking by most OECD countries (OECD, 2020_[23]). This is true also inside MIMIT, where there is a consistent need of technical skills and data infrastructures. For this reason, in order to produce evaluations, it is essential for MIMIT to invest in people and in data and OECD countries offer some good practices.

To produce evaluations, it is essential to have the right skills inside the Unit. MIMIT's Centro Studi faces two problems in this context: 1) it lacks, in its set-up stage, the sufficient number of experts to perform policy evaluations and 2) senior civil servants with managerial responsibilities (i.e., "Dirigenti" and above) often have a legalistic culture, rather than an evaluation culture. Some remedies adopted from other OECD countries to address these challenges were to create specific recruitment tracks and career paths for employees with analytical skills within the civil service and investing in trainings both for evaluators and decision-makers. Particularly, recruitment tracks and career paths are considerably different from the current Italian way of recruiting, which is based on general public exams.

BOX 3.1: GOOD PRACTICES ON EVALUATION PLANS

The Swedish Agency for Growth Policy Analysis, produces every year a three-year [evaluation plan](#) which states the evaluations that they will undergo in the following three years and sets a delivery date. This plan is updated annually to ensure that it is in line with shifting political priorities. As it emerges from their evaluation plan, they perform analysis both on broader economic questions (effects of AI on enterprises, digitalisation and internationalisation) but also policy-focused evaluation like the Evaluation of the R&D deduction which will be published in September 2022.

Similarly, the Research Department of VisitINPS also grounds its evaluations on a three-year plan which is updated every year. This is approved by the Director of the Department, together with the Board of INPS to align with INPS strategic interests and priorities. The evaluation plan is a chance to engage with policymakers and ensure relevance of the research conceptions. This is slightly different from the approach adopted in the Netherlands, where the Bureau for Economic Policy Analysis (CPB) sets workplans in collaboration with the ministry but is the ultimate independent decision-makers on the topics that it considers most interesting and suited for evaluations and the methodologies and approaches to conduct these evaluations.

The adoption of an evaluation plan is also standard practice at the US Department of Labour (DOL), which, following the Evidence Act of 2018, is required to publicly describe future plans and current capacity to conduct evaluations and other evidence-building activities across three planning reports: the Annual Evaluation Plan, an Evidence-Building Plan, and a Capacity Assessment for Research, Evaluation, Statistics, and Analysis.

Source: (Tillväxtanalys, 2022^[28]).

To perform evaluations, it is important to hire people with specific skills and backgrounds. These are not only statisticians, data scientists and economists, but also policy analysts and social scientists. Even if multiple backgrounds can enrich policy evaluations, the skills that should be prioritised are the analytical and data skills. In some cases, especially at the beginning, these might be limited and for this reason some evaluations could be out-sourced, e.g. to other public institutions and research institutions (See Box 3.2).

BOX 3.2: OUTSOURCING EVALUATIONS: THE BMWK MODEL

Evaluations can require significant expertise and resources. In some cases, these resources might not be available inside the Ministry and commissioning external evaluations represents a way to deal with this problem. This can be a good option especially at the beginning of the Centro Studi's activities where the number of employees is still limited. This practice is used by several OECD countries. For example, both the BEIS in the UK and the BMWK in Germany collaborate with external parties.

The BMWK regularly publishes the evaluation they perform and these come often from external actors. Indeed, BMWK supports eight economic institutions (DIW, ifo, IfW, IWH, LIF-SAFE, RWI, ZBW, ZEW) and the Leibniz Institute for Applied Geophysics (LIAG) with the aim of obtaining excellent and up-to-date research and evaluations. In 2021, 51 research and evaluation contracts were awarded with funds amounting to around 7.6 million. In addition, research assignments are also financed from specialist titles of the other departments. Instead of evaluations, the internal staff in some departments of the Ministry prepare analysis and studies on economic issues which are published as Discussion papers.

Despite not producing evaluations in-house, the Ministry remains a fundamental stakeholder and participates to meetings and reviews. Moreover, the BMWK is equipped with a Scientific Advisory Board that has an outstanding reputation for scientific policy advice, being the first independent body of scientific policy advice in post-war Germany. These are currently 41 scientists which are appointed by the Federal Minister of Economics and Climate Protection. This group meet five times a year and publish expert opinions.

Source: (BMWK, n.d.^[29])

BOX 3.3: PROFESSIONALISATION OF POLICY EVALUATION IN IRELAND, THE UNITED KINGDOM AND CANADA

The **Irish Government Economic and Evaluation Service** (IGEES) was created in 2012 to develop analytical capacities for evidence-informed policy making across the Irish Government. The IGEES staff is located in each department to offer specific analytic and expertise. This staff is composed of graduates, experienced economists, evaluators and policy analysts. This was possible through directly recruiting with an open competition process. Moreover, IGEES offers also opportunities for capacity building and skills enhancement through structured mobility, a learning and development framework and targeted opportunities.

The **UK Government Social Research** (GSR) Profession represents civil service professions that works alongside economists, statisticians and operational researchers. These civil servants use social research methods (surveys, controlled trials, qualitative research, cases studies, analysis of administrative and statistical data) to understand and predict relevant phenomena for policymaking (OECD, 2020_[23]). The GSR profession has a specific competency framework that begins with entry-level graduates as part of the fast stream to members of the senior civil service. Most UK government departments have a chief social researcher who leads and supports the activity of social researchers within the department.

In Canada, the **Treasury Board of Canada Secretariat** (TBS) adopted a list of evaluation competencies that each evaluator at different level must possess. This is organised in terms of evaluation skills, integrity and respect, vision and strategy, collaboration etc. It provides a baseline common to all evaluators that ensures quality and harmonises the evaluators' skills.

Source: (OECD, 2020_[23])

BOX 3.4: TRAININGS TO DEVELOP POLICY-MAKERS SKILLS TO USE EVIDENCE, UK AND PORTUGAL

In the UK, the Alliance for Useful Evidence holds a masterclass where policy makers learn about how to use evidence and have the possibility to directly practice the skills learnt through simulations (OECD, 2020_[23]). Moreover, the United Kingdom's Government Digital Service (GDS) Academy¹ provides training for public servants to enhance digital skills. Courses focus on the benefits and challenges of the digital transformation, and the relevant soft and hard skills. Courses focus on both managerial aspects related to the use of data as well as the technicalities of applying data to the design and delivery of public services. This mix of competencies results indeed critical to have a shared and data-drive approach within and across public sector organisations (Van Ooijen, 2019_[6]).

In Canada, in the field of healthcare, there is a specific executive training in research application (EXTRA). This provides support to policy makers to use evidence and to train their co-workers.

In Portugal, every year, public service organisations inform the National Institute of Public Administration of the needs of their employees. This allows for the creation of annual training programmes able to address the specific needs of public administrations.

Australia is also a good example in terms of efforts made to boost data analytics skills for public servants. The Australian Public Service (APS) "Data Skills and Capability Framework" is based on four pillars: data fellowship, university courses, training partnerships, and a data literacy programs. On the one hand, the data fellowship program and training partnerships enable public servants to join private sector organisations to work on data-related problems. On the other, university courses and data literacy programs allow to strengthen and update data skills with the most recent advancements in the field.

Source: (OECD, 2020_[30]) (OECD, 2020_[23])

1. <https://www.gov.uk/guidance/gds-academy-courses>

PRODUCING HIGH-QUALITY RESULTS

Quality of evaluations is essential to ensure a positive impact on policy making (OECD, 2020^[23]). Nonetheless, achieving quality represents one of the main challenges that countries face when implementing evidence-based policymaking. This will be true also for the Centro Studi, especially in its initial phase. To ensure that evaluations will be of good quality, several mechanisms are in place in other OECD countries. These can be divided in Quality control (through guidelines and evaluation of the end product) and quality assurance processes (through promoting the quality at the institutional level and of evaluators).

The use of guidelines, peer reviews groups and scientific councils are ways to ensure quality. According to the 2018 OECD Survey on Policy evaluation, 29 of 42 respondent, of which 24 OECD countries, have adopted instruments to boost technical quality and good governance of the evaluation process. There are four main ways that were identified to increase quality of evaluations (OECD, 2020^[23]):

- 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 end product;
- Supporting and promoting evaluator competences;
- Fostering quality at an institutional level.

MIMIT might benefit from developing some quality standards through specific evaluation guidelines. Indeed, even if some general evaluation guidelines are already established by the Presidency of the Council, the Centro Studi could complement them with additional ones more specific for the MIMIT policies, building on the activities of this project. Indeed, in several OECD countries it is common to have guidelines at different levels. For example, the United Kingdom is equipped with several guidelines to carry out evaluations (See Box 3.5). Some of them are established at the central level, like the Green Book and the Magenta Book. However, departments such as the BEIS, which is close to MIMIT in terms of portfolio, can produce internal guidance specific to their areas of responsibility. BEIS regularly produces more specific guidelines for quality assurance (QA) like the [BEIS Quality Assurance \(QA\) modelling: guidance for models](#). This internal guidance should be aligned with the governmental guidelines to guarantee consistency.

BOX 3.5: EVALUATION GUIDELINES IN THE UNITED KINGDOM

The Green Book: Central Government Guidance on Appraisal and Evaluation. Firstly published in 2003, it has been updated in 2011 and 2018 by the HM Treasury. It provides guidance to officials to develop transparent, objective and evidence-based policy evaluations.

The Magenta Book: Central Government Guidance of Evaluations. It was published in 2011 and updated in 2020. The objective is to provide an overview of evaluation in government. It provides guidance on policy evaluation and presents also different type of evaluation methods (process, impact, value-for money).

The Aqua Book: Guidance on producing quality analysis for government. Its objective is to improve the analytical quality assurance. It helps the department and agencies to embed an analytical environment that assists the delivery of quality analysis (OECD, 2020^[23]).

BEIS Quality Assurance (QA) modelling: guidance for models. Its main objective is to improve the quality of the models that are created or used inside BEIS. This is not specifically focused on policy evaluation, but modelling represents a very widespread evaluation method inside BEIS and for this reason provides some interesting technical information.

Source: (OECD, 2020^[23])

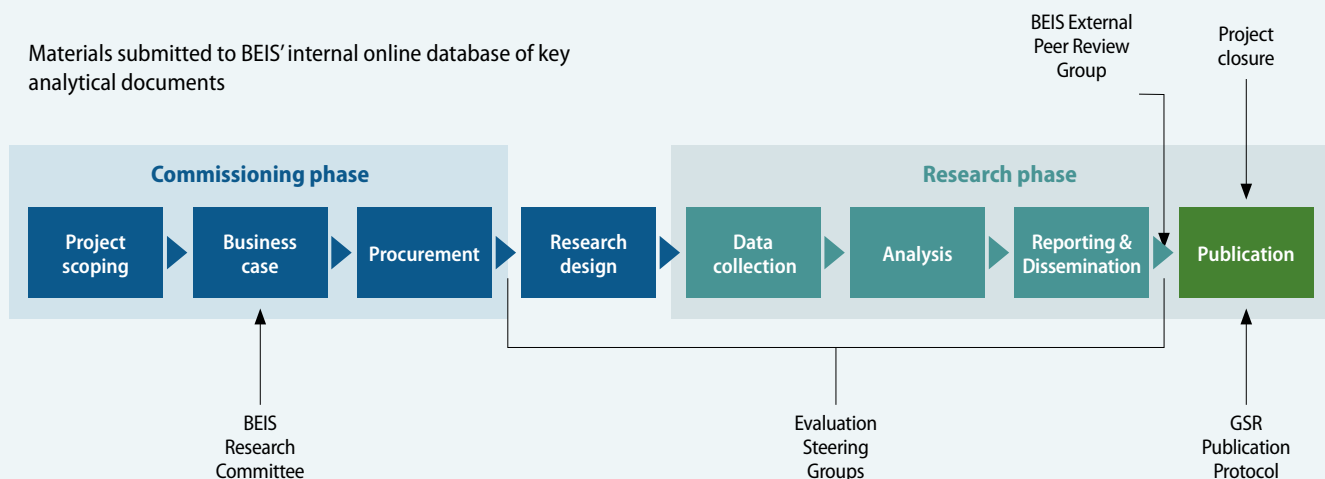
Once the evaluation process is completed and the product is finalised, it is important to have mechanisms in place to ensure the quality of the final product. To do so, external peer review networks can play an important role. This method is used by several countries like Germany, the UK and Canada. These networks are fundamental to safeguard the impartiality of the findings and increase accountability. In particular, the UK BEIS uses a series of both internal and external peer reviewers to foster quality of the evaluation (See Box 3.6). Another control mechanism is to establish a Scientific Council composed by experts responsible of reviewing the work inside the evaluation unit. This method is used by the Growth Analysis unit in Sweden and also by the VisitINPS programme in Italy (See Box 3.7).

BOX 3.6: PEER REVIEW GROUP (PRG) IN THE UK DEPARTMENT OF BUSINESS AND ENTERPRISES (BEIS)

To ensure independent, transparent and high-quality evaluations, BEIS uses a **Peer Review Group (PRG)**. This represents an external group of experts (generally two peer reviewers with expertise in the field) which are consulted at key stages of the evaluation. All evaluations need to undergo this independent peer review and comments need to be addressed by internal evaluators (See Figure 3.1). Moreover, when not all the comments suggested by the PRG are addressed in the evaluation, these are published in order to make them accessible and visible to the public.

FIGURE 3.1. UK BEIS evaluation governance

Materials submitted to BEIS' internal online database of key analytical documents



Source: (Department of Business, Energy & Industrial Strategy, 2020_[31])

BOX 3.7: SCIENTIFIC COUNCIL OF GROWTH ANALYSIS IN SWEDEN

The **Swedish Agency for Growth Policy Analysis**, is an agency that evaluates growth policies in Sweden. The agency works by government commission and is supervised by the Ministry of Enterprise and Innovation, but it also collaborates with other ministries in the area of growth policies. It employs 45 people. The agency main task is to evaluate the policies requested by the government. In addition to this, it has also some space for self-initiated projects. In order to support the researchers inside the Growth Analysis, the government appoints a Scientific Council. This Scientific Council has the task of strengthening the unit with methodological support. The work of the Council is led by the Director General. In this moment, the Council is composed of five academic experts coming from different Swedish universities.

Source: (Tillväxtanalys, n.d._[32])

The quality of the evaluations is strictly related to the quality of the evaluators. To ensure the quality of the evaluators, it is important to be able to hire evaluators and analysts with the right skills as previously discussed. Training programmes and informal networks can increase the quality of evaluators and ensure the promotion of an evaluation culture. A good example is again BEIS which engages several internal and external networks in order to promote knowledge sharing (See Box 3.8).

BOX 3.8: UK BEIS'S EVALUATION NETWORKS

BEIS takes part in several networks in order to increase the analytical capacity and capability. These networks are both internal and cross-governmental:

The **Evaluation Practitioners Network (EPN)** is an internal network that meets monthly to support the delivery of good quality evaluation in BEIS.

The **Cross-Government Evaluation Group (CGEG)** aims to improve supply, demand and use of policy evaluations inside the government. It is a cross-departmental and cross-disciplinary group.

The **Cross-Government Monitoring, Evaluation and Learning Official Development Assistance (ODA)** Group is a knowledge exchange with monitoring and evaluation advisers across government working on ODA.

Source: (Department of Business, Energy & Industrial Strategy, 2020^[31])

COMMUNICATING RESULTS IN AN EFFECTIVE WAY

Communication is essential to ensure that evaluations, once produced, are also used and known both inside the Ministry and outside of it. Other OECD countries have implemented a series of practices to ensure that communication is effective and evaluations are diffused, such as the use of portals or the creation of tailored-products that can be more accessible to decision-makers.

Promoting access to evaluations is a way to ensure that policy-makers will be informed about the results of the evaluations, while increasing transparency and accountability. In several OECD countries, evaluation portals are used to make evaluations easily available. The Norwegian evaluation portal,¹ is a publicly accessible web service that gathers all the findings of evaluations carried out by the central government. Examples also exist at the level of single line ministries. In Canada, the ISED provides an evaluation page² where all past evaluations are collected in order to make access easy.

Ensuring that information is publicly available is not enough to ensure impact. To increase the chances of uptake it is important to provide tailored products, like policy briefs or summaries of findings. Some research methods can also help collecting different dispersed knowledge and create final products which summarise evidence. Systematic reviews and Rapid Evidence Assessment represent interesting methods. They are widely used in the realm of health and becoming increasingly popular in social sciences. As the time of research and of political decision are often very different, creating these products represents a way to facilitate policy-makers' access to evidence.

1. Accessible at: <https://evalueringsportalen.no/>

2. Accessible at: <https://ised-isde.canada.ca/site/audits-evaluations/en/evaluation>

4. Enhancing the Centro Studi's data capacities drawing on international best practices

This section reviews relevant best practices on data collection, integration, analysis, storage, and visualisation methods, currently used in other countries and/or national institutions that could provide inspiration to MIMIT to address the challenges discussed in Chapter 1.

Countries differ significantly in their policies and organisational set-ups for data collection, integration, storage and sharing – partly due to differences in privacy laws and governance structures. Nonetheless, data architectures and infrastructures enabling effective data management and policy evaluations have usually common features, and several lessons and insights of relevance to the Ministry can be drawn from international practices.

Traditionally, data management in public administrations suffers from a ‘silos’ approach. Indeed, formal data management is often conceived as a localised function, under the responsibility of technicians and information and/or data managers (Ladley, 2019^[33]), resulting in fragmented and localised information (Ubaldi, 2013^[34]). These data silos create challenges in fostering data interoperability, creating accessible registers, and supporting data sharing. In addition to missed opportunities/synergies and a less efficient use of public resources, data silos hinder opportunities for data-driven policy-making, as data integration is a key prerequisite to perform comprehensive policy assessments.

To transform raw data into policy-relevant information and analysis, governments need to effectively collect, integrate, process and share data. The “data value cycle” is a useful conceptual framework to understand the relationships between these activities (Van Ooijen, 2019^[35]). As highlighted in Figure 4.1, one can identify four main steps for an effective data exploitation. First, data from governmental and non-governmental sources have to be *collected and integrated*, through the use of, e.g., administrative information, surveys, and relevant open data. Second, data have to be *processed, secured and stored*, in order to valorise collected information, prevent data silos, and ensure quality and security. Third, relevant capabilities should enable data to be *curated, shared and published* within and outside the public sector, key elements for allowing, fourth, an *effective use of data* for policy analysis and policy making. Indeed, data sharing, access, visualisation and analysis have the potential, on the one hand, to directly increase public value – since collected information and data can be directly employed to ameliorate policies and services. On the other, an effective use of data induces positive feedback loops on the creation and demand of high-quality data from governmental and non-governmental sources, increasing the potential of the data value cycle itself. All in all, data which flow through a structured value cycle of data collection, integration, management and sharing/analysis have the potential to positively impact public stakeholders and increase public value.

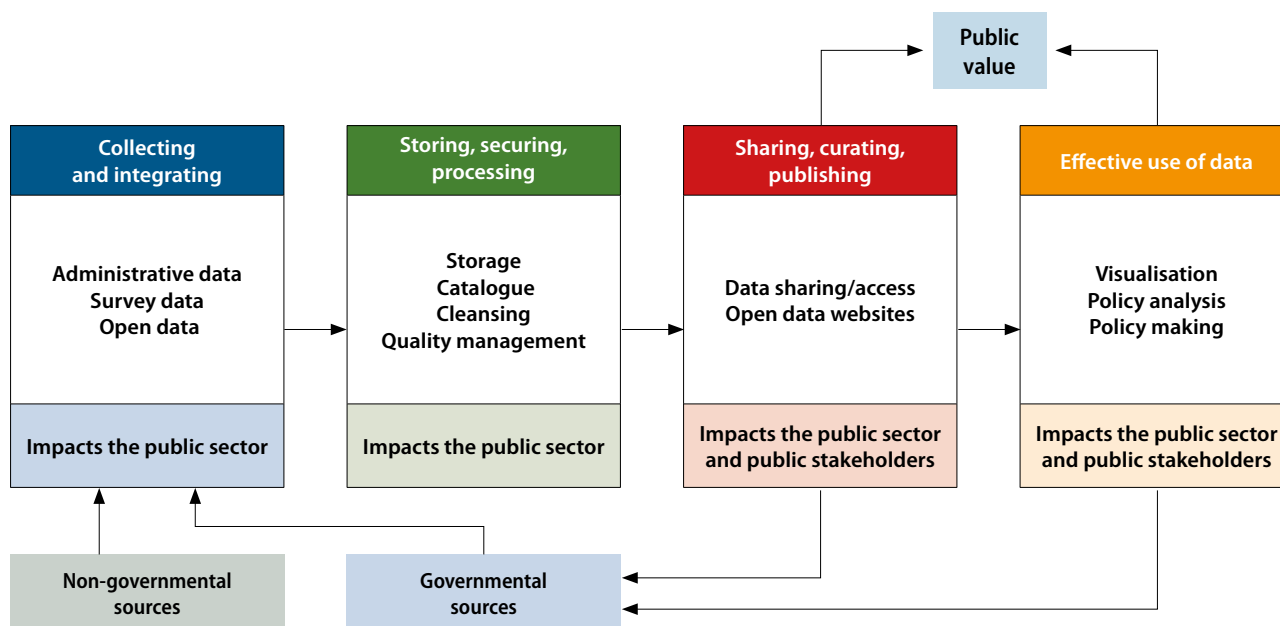
This section reviews each aspect of the data value cycle, focusing in particular on the best practices to effectively collect, integrate and harmonise data sources (both administrative and survey data), the cutting-edge solutions for data storage and management, and the most suitable methods to allow researchers to access confidential microdata to conduct policy analysis and research. Finally, the section presents some of the most effective visualisation tools that have been developed to support policymaking.

DATA COLLECTION: ADMINISTRATIVE AND SURVEY DATA

Data are the backbone of evidence-based policy. Yet, obtaining the data, extracting relevant information, and integrate different data sources are usually the most time-consuming and intensive parts of policy monitoring and evaluations.

As showed in the data value cycle (Figure 4.1), **data collection** (and generation) from governmental and non-governmental sources is the first fundamental step for creating public value through the use of data. To timely inform policymaking, it is crucial that microdata are available and integrated across different sources and that **data collection is planned** at the beginning of the policy cycle.

FIGURE 4.1. The data value cycle



Source: elaboration based on Van Ooijen (2019_[35]).

A key distinction can be made between administrative and survey data sources:

The use of **administrative firm-level data** for policy analysis has flourished over the last two decades. Their characteristics –being regularly updated, consistent, and with a wide coverage of the population of reference– has considerably boosted the capability to perform evidence-based policy (Box 4.1). Several countries have made substantial investments in administrative data collection – as well as integration and sharing. These countries include, among others, **New Zealand, UK, Germany, Denmark, France, Canada and the Netherlands**, as it will be discussed more in details in the next sections. These achievements were made possible through extensive efforts to foster relationships with data providers and to improve data-sharing agreements, ultimately enabling better data usage for evaluations. In Italy, ISTAT has considerably strengthened its ability to collect and integrate administrative data and widened the set of micro-data providers.

The use of administrative data to support evidence-based policy has also increased thanks to the **easing of their accessibility** for research purposes. Indeed, several institutions and Statistical offices now provide access to confidential administrative data, both within the public sector (e.g., to Ministries) and to public stakeholders (e.g., academia, through research projects – as discussed more in detailed in the next sections).

Survey data provide a crucial input for policy making. Surveys may provide more **timely information** than administrative data, thus enhancing the monitoring and evaluation of policies (since, as previously discussed, administrative data are usually available with one/two years lag). At the same time, they can be used to complement administrative data, by **eliciting additional important information** (such as firm's expectations about demand and supply, as well as knowledge-based, organisational and managerial assets). To effectively support evidence-based policy, surveys need to be properly designed. Box 4.2 provides guidance to this end.

For example, focusing on firms, the UK Department of Business, Energy and Industrial Strategy administers the yearly **Longitudinal Small Business Survey**. The LSBS has been collecting data on a large sample of mSMEs since 2007. Its long-term perspective and the consistency of some key information overtime has allowed LSBS to be widely used in both economic papers and policy notes.

BOX 4.1: THE USE OF ADMINISTRATIVE DATA TO SUPPORT EVIDENCE-BASED POLICY MAKING

Administrative data have been recognised as powerful tools to support evidence-based policymaking. While administrative data are not collected for research – but to meet administrative requirements – their characteristics are particularly suitable for evaluation purposes.

First, since these data are already collected by government and other agencies – while conducting their activities and providing services – the use of administrative data is cost-effective and cheaper than collecting the data *ex-novo*, e.g. through surveys (Crato and Paruolo, 2019^[36]). Moreover, administrative data provide a large set of information covering a large sample size that is usually representative of the whole population of reference (e.g., businesses affected by a specific program or regulation). Administrative data are usually regularly updated, and the same unit of analysis observed repeatedly over time. This *longitudinal* component of data is crucial to go beyond a single follow-up of the measure and evaluate policies more efficiently than e.g., using one-time survey.

However, administrative data are also subject to some limitations (Crato and Paruolo, 2019^[36]). For example, despite being regularly updated, there might be delays in the data collection process. Moreover, confidentiality issues may prevent data to be shared or likable with other data sources. Finally, administrative data may not include relevant-specific information that are crucial for policy evaluations. To address some of these issues, administrative data can be integrated with other data sources that may complement and enrich their information, such as survey data – that will be discussed below.

BOX 4.2: DESIGNING SURVEYS TO SUPPORT EVIDENCE-BASED POLICY

Surveys have the potential to increase the timeliness and effectiveness of policy evaluation. Yet, this potential may not be fully exploited if surveys are not properly designed.

Information collected: if surveys are designed simply to collect information, e.g., on policy take-up, the information collected may be limited to the use of the policy, possibly with some overall evaluation by the firm which is akin to a sort of “customer-satisfaction” question. Surveys should be designed to collect broader information e.g., on firm performance, management practices, expectations etc.

Control group: surveys should not limit the collection of information to beneficiaries of the incentive. Sampling non-beneficiary firms allows to create a useful control group for performing counterfactual policy evaluations.

Follow-ups: surveys are not always accompanied by subsequent follow-ups, limiting the possibility of following the performance of beneficiaries over the years. Even when surveys are repeatedly conducted, it might occur that the data are not collected for the same units, preventing the use of a longitudinal/panel data analysis.

Quality checks: surveys can be subject to measurement errors (due to sampling errors, non-responses, representativeness issues etc). Checks and assessments are necessary to ensure the quality of collected data.

Matching with administrative data: surveys should be designed to allow matching with other data sources, such as administrative data, and questions should be crafted to elicit complementary information. One example of survey specifically designed to be match with other data is the COVID-19 survey administered by ISTAT to monitor firms' economic performance and digitalisation during the pandemic. The survey was administered to a subset of firms that participated to the 2018 Census, while maintaining national representativeness. This feature was exploited by Calvino et al. (2022^[37]) to conduct a rich analysis on digitalisation of Italian firms during the COVID-19 crisis.

In Italy, the **INVIND Survey** conducted by the Bank of Italy on Italian business represents an extremely successful case (see Chapter 2), while the **RIL Survey** (Rilevazione Longitudinale su Imprese e Lavoro) administered by INAPP has also been widely used for policy evaluation.

In general, effective data collection and generation must be complemented by the collection of detailed and comprehensive information about the data themselves (**metadata**). Metadata allow data to be re-usable, exchangeable and preserved (Bender, Hausstein and Hirsch, 2018^[38]). Each data source should include appropriate information on how data have been collected, how they are organised, which variables are included, and who can access them.

DATA INTEGRATION, HARMONISATION, AND INTEROPERABILITY

Integrating different data sources is **crucial to perform policy analysis and evaluations**. Indeed, integrating microdata allows to exploit more comprehensive and disaggregated information that are key to better evaluate firms' response to policy measures. Moreover, the integration of data sources facilitates research into different policy areas for which policy silos may limit their connection (Langedijk, Vollbracht and Paruolo, 2019^[39]).

However, **integrating data is usually costly and time-consuming**. On one hand, data linkage raises privacy and confidentiality issues: the need for data anonymisation restricts data access and content and might prevent data to be integrated. On the other hand, when data integration is possible, it presents several methodological issues that may reduce the effectiveness of the integration. First, datasets must contain information enabling records to be associated – directly through a unique identifier or indirectly using matching techniques.¹ Second, information contained in different datasets might not precisely coincide, inducing consistency issues. For example, differences across datasets may occur in the unit of analysis (e.g., firms or establishments), in the definition of variables of interests, and in the data collection process (e.g. timeliness). Moreover, the classification used in one data source may not be the same as in other datasets.

Several international best practices have successfully overcome some of these challenges, creating integrated databases linking administrative and survey data. Among them, it is worth mentioning the New Zealand integrated data infrastructure including **the Longitudinal Business Database (LBD)** and **the Integrate Data infrastructure (IDI)** (see Box 4.3), the **Institute for Employment Research (IAB)** of the German Federal Employment Agency (discussed in Box 4.7.), and the **Administrative data research Network** in UK (ADR-UK).²

Another good example of data integration is provided by the **ESANE system of INSEE** (the French National Statistical Office). ESANE builds a comprehensive database of enterprises by integrating administrative (Unified System of Business Statistics, SUSE) and survey data (annual business survey, EAE)³ to produce accurate structural business statistics. In Colombia, the Base Statistical Registry of Companies (REBE) integrates information from Chambers of Commerce, social security, tax authority, inspections of companies, and surveys conducted by the statistical office DANE.

Better data integration can be achieved when different sources of data are **harmonised** ex-ante, as a result of well-conceived **data architectures**. In this sense, *shared standards, consistent logic across data, the use of uniform technical formats, and appropriate data access rules* are key elements of effective data architectures, ultimately enabling better **data interoperability** (as demonstrated, e.g., by **Estonia** – see Box 4.4).⁴ For example, data have a consistent logic when information about data (metadata) are shared, and uniform semantic conceptions render the combined use of data possible and effective. Semantic interoperability means avoiding the collection of data following different semantic conceptions for the same information/variables (e.g., information about investments made by companies), and having shared standards that ensure the unambiguous meaning of information (Domeyer et al., 2021^[42]).

1. For more details on data linking techniques see Shlomo (2020^[40]).

2. The ADR-UK provides integrate data from different UK governments agencies, thanks to a partnership between four Administrative Data Research Centres (England, Northern Ireland, Scotland, Wales) and the Office of National Statistics (ONS). The project is coordinate by a UK-wide Strategic Hub, and it was founded by the Economic and Social Research Council. These data are also made accessible to researchers. See <https://www.adruk.org/about-us/about-adr-uk/> for more details.

3. See <https://www.insee.fr/fr/metadonnees/source/serie/s1188>

4. For more details about the key elements of data interoperability see (Domeyer et al., 2021^[42])

BOX 4.3: INTEGRATING AROUND A SPINE: THE EXPERIENCE OF STATISTICS NEW ZEALAND

Stats NZ has two rich integrated databases: the Integrated Data Infrastructure (IDI) and the Longitudinal Business Database (LBD). The IDI contains de-identified data on individuals, while the LBD contains integrated data on firms. Both databases combine several data from Stats NZ surveys and government agencies.

In both cases, data integration is built around a central dataset, called the “spine”, to which all other data are linked to (Figure 4.2). The IDI “spine” aims to include all people living in New Zealand and was created using probabilistic linkage (Shlomo, 2020^[40]), linking tax data to births data, births to visa data, and visa to tax data, and combining them. This data is then linked to information on education, health, social welfare, income and justice.

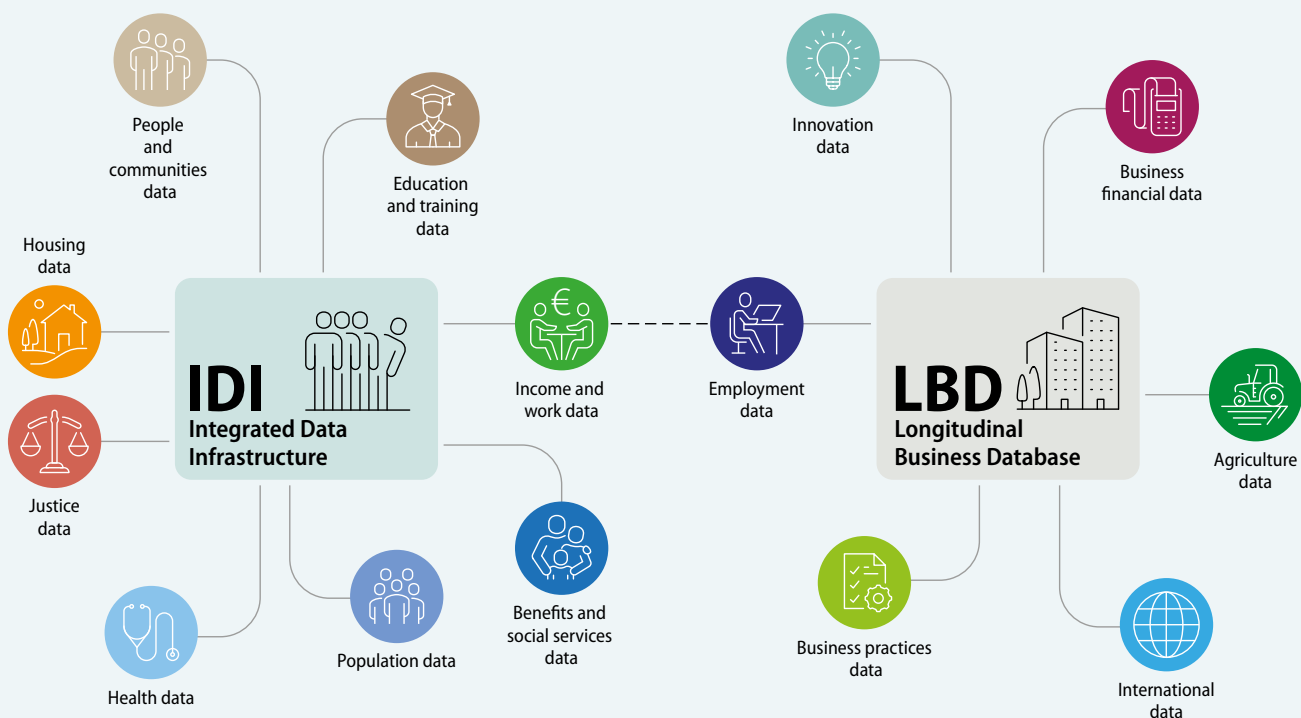
The LBD has its backbone in the Longitudinal Business Frame – a longitudinal register of business in New Zealand – to which other firm-level data are linked, such as data on trade, agriculture, business practices, financial and firms’ innovation.

LBD and IDI can be linked together via employee tax information to study the longitudinal linkages between individuals’ characteristics, their employment history, and firms’ activity.

The databases are updated regularly (e.g., every year for the LBD), and the previous versions are stored and accessible when required. One important feature of IDI database is that it is continuously improved based on the administrators’ and users’ experiences.

Access to IDI and LBD is provided to external researchers, subject to specific requirements summarised in the “Five Safes” principles—see below for more details. While Stats NZ integrates the datasets using identifiable data, it then removes personal information when data is made accessible to researchers to ensure no association of records to specific individuals is possible.

FIGURE 4.2. Integrated databases in Stats NZ on individuals and business: Integrated Data Infrastructure (IDI) and Longitudinal Business Database (LBD)



Source: (Jones et al., 2022[41]), <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/>.

BOX 4.4: BEST PRACTICES ON DATA INTEROPERABILITY

Estonia has been one of the first countries that achieved **effective interoperability of data collected and managed by different institutions and agencies**. The use of unique identifiers for citizens and companies and the integration and systematisation of data sources from different institutional bodies have enabled Estonia to link relevant registers and make them interoperable. Today, Estonia operates its registers through the X-Road data exchange system, an open-source data exchange solution that also allows secure interoperability among Estonian and Finnish national population registers (following the connection of business registers and tax boards in 2019). This has resulted also from the approval, in 2000, of the “Public Information Act”, which prohibited the establishment of separate databases for collecting the same data and enhanced the connected and consistent use of available information (“once-only” principle). In order to achieve effective data interoperability, the once-only principle has been implemented across agencies and ministries, in order to achieve more efficient data systematisation primarily within relevant levels of the public sector.

Note: The X-Road data exchange mechanism is managed by the “Nordic Institute for Interoperability Solutions” (NIIS), a non-profit association with the mission to ensure the development and strategic management of the X-Road data exchange system and other solutions for digital government infrastructure. The republics of Estonia, Finland and Iceland are members of the NIIS.

Source: (OECD, 2019^[43]; X-Road, 2022^[44]; NIIS, 2022^[45]); <https://joinup.ec.europa.eu/collection/ict-security/solution/x-road-data-exchange-layer> <https://www.riigiteataja.ee/en/eli/514112013001/consolide>

IMPROVING DATA CAPABILITIES: DATA INFRASTRUCTURE AND SKILLS

To effectively **employ collected data**, break down data silos, and foster a data-driven policy making, it is critical to promote better **data management practices** within government agencies. Data management deals with the **processing, computing and storing** phase of the data value cycle (Figure 4.1), that is critical to enable effective down-stream data access, sharing and analyses (discussed in the next section), as well as to positively impact the public sector and other public stakeholders (Van Ooijen, 2019^[35]). In other words, better data management critically relates to the design of effective data architectures and investments in data **IT infrastructures** and **skills**.

Data infrastructure refers to an ensemble of technical components (ICT investments) and decisions for **storing, securing and managing data**. The precise technical choices must necessarily adapt to the specificities of single organisations and respond to public procurement constraints and objectives. Available broad solutions for data storage relate for example to the use of in-house and/or cloud-based servers (closely linked to the issue of security), while data management depends both on the architecture of data and related data management systems.

Data warehouses and **data lakes** constitute two prominent solutions of data management systems within public and private organisations. These two exemplify the challenges and trade-offs that data management choices imply for organisations, both in the short and long term (see Box 4.5). For instance, while having a relatively simpler data management system (e.g., data warehouse) may be useful for internal rapid data visualisation and analysis, a more comprehensive set of information (as in data lakes) would be potentially useful to improve data usage and exploitation, also for external stakeholders. At the same time, while ingesting more information in a data management system enables – in principle – wider and richer analyses, increased complexity requires appropriate ICT and data skills to fully exploit available information.

Having a **centralised data management system** may enable better interoperability and transparency, as well as data collection and integration **automation**. Indeed, localised data management practices induce the collection of fragmented information, the duplication of costs and, eventually, reduce productivity and efficiency (Ubaldi, 2013^[34]). On the contrary, the centralised management of data enables economies of scale in collecting, systematising, and analysing relevant information, both for monitoring and evaluation purposes. In this sense, the effectiveness of data management system is tightly related to storage solutions and investments in the **IT infrastructure**, enabling the centralised systematisation of data.

BOX 4.5: DATA WAREHOUSE AND DATA LAKE: DATA MANAGEMENT SYSTEMS IN THE ERA OF BIG DATA

A **data warehouse** is a system that **aggregates structured data from different sources into a centralised data store**, which then connect those data to analytical tools to perform data analyses.

Data warehouses follow a “schema” data model, where data must be structured using predefined schema. This is usually accomplished by a three-tier architecture. The bottom tier consists of a data warehouse server which collects, cleanses, and transforms data from multiple data sources through an “Extract, Transform, and Load” (ETL) process. Defining schema requires knowledge and planning of which data will be used, since data structure has to be optimised and reduced (to fit schema) before entering the warehouse. The middle and top tiers consist respectively of an OLAP (i.e., online analytical processing) server, which allows query speeds, and by front-end user interface (e.g. reporting tool), which enables end users to conduct ad-hoc data analysis and visualisation on their data.

A **data lake** is a centralised repository where structured, semi-structured, and unstructured data can be stored in raw origin format, from different sources. This “schema-on-read” process enables the aggregation and transformation of data at query-time, bypassing the ETL transformation process typical of data warehouses (where data are elaborated ex-ante to fit pre-defined schema). In this sense, **a data lake can be interpreted as a data warehouse without predefined schema, where inputs do not have to be processed in advance.**

Table 4.1 summarises strengths and weaknesses of the two systems.

TABLE 4.1. **Data warehouse vis-à-vis data lake: strengths and weaknesses**

	Strengths	Weaknesses
Data warehouse	<ul style="list-style-type: none"> • Allows better ex-ante data quality (ETL cleanses, eliminates duplicates, and standardises data). • Enables data integration from different sources. • Supports large-scale functions such as data mining, artificial intelligence, and machine learning in an easier way (finding unseen patterns and relationships in data) 	<ul style="list-style-type: none"> • Data without an ex-ante clear use case are often discarded. • Ingesting more data requires high computational resources for the ETL process.
Data lake	<ul style="list-style-type: none"> • Helps eliminate data silos, since it is a single landing zone for data from multiple sources. • Ingests large volumes of data at a low computational cost, including data without a clearly defined use case. • Promotes agility and enables to query data in novel ways and uncover new use cases 	<ul style="list-style-type: none"> • Does not primarily include analytic features. • Without governance, integration with analytics tools, and appropriate human capital (ICT) skills, it can become a “data swamp”, where data accumulate without effective use.

Source: (IBM, 2020_[46])

Having a **modern and up-to-date IT data infrastructure** is nowadays critical for public sector organisations, for the productivity and costs benefits coming from better data management and exploitation, as well as for reducing data and confidentiality risks.

The choice of **data storage** solutions links to data access, sharing, and security, and broadly relates to technical decisions among, e.g., **in-house** and/or **Cloud-based servers**. In-house servers and Cloud-based services constitute two major solutions in the field of IT data infrastructure (see Box 4.6), which can be combined to promote public sector digitalisation (World Bank, 2022_[47]). **Hybrid cloud-based solutions** that mix public and private cloud services for data storage and processing has brought great benefits for public agencies and stakeholders, both in terms of data accessibility and security (Jones et al., 2019_[48]). Indeed, while allowing for wider data access vis-à-vis locally installed servers, hybrid cloud solutions allow public sector institutions to maintain direct control on sensible data and limit potential security issues (Van Ooijen, 2019_[35]).

BOX 4.6: DATA STORAGE: IN-HOUSE SERVERS AND CLOUD-BASED SOLUTIONS

In-house server solutions leverage LANs (local area networks) technologies to store data in-house, within a local datacentre. Dedicated servers rely on physical IT data infrastructure and capacity, and require **ICT skills for constant upgrading and maintenance**, as well as frequent **IT investments** in up-to-date technologies. Local servers give organisations direct control over data and support security (if data are managed centrally and the number of access points reduced).

Cloud-based solutions leverage instead internet-based servers to store data, which enable, in general, increased levels of efficiency and **cost reduction for managing the physical infrastructure**, as well as greater workforce adaptability for remotely accessing information.

TABLE 4.2. **In-house servers vis-à-vis Cloud-based solutions: strengths and weaknesses**

	Strengths	Weaknesses
In-house servers	<ul style="list-style-type: none"> Better security in presence of centralised servers Higher control on confidential data Higher internal accessibility and flexibility for managing data 	<ul style="list-style-type: none"> Usually require higher capital investments and maintenance costs Lower degree of data sharing and interoperability with external stakeholders
Cloud-based solutions	<ul style="list-style-type: none"> Reduce maintenance costs related to ICT infrastructure Increase workforce adaptability and remote working solutions 	<ul style="list-style-type: none"> More binding storage limits More prone to data-security issues and risks of cyber attacks

Cloud computing offers various service models such as software, platform and infrastructure services (OECD, 2014_[49]). These services can be categorised into **“software as a service” (SaaS)**, **“platform as a service” (PaaS)** and **“infrastructure as a service” (IaaS)**.

IaaS provides computing resources such as processing, storage and networks to the users of clouds, enabling usage of own virtualisation capabilities (leading to maximum flexibility). In the **SaaS** solution, there is direct access to applications of the cloud provider, without the need of managing the underlying infrastructure or applications, with included tools for specific analytical

DATA SHARING: GRANTING MICRO-DATA ACCESS FOR RESEARCH AND BUILDING STRONG COLLABORATIONS WITH ACADEMIA AND RESEARCH CENTRES

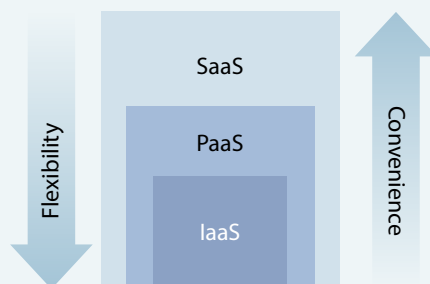
Granting micro-data access to external researchers is key to boost high-quality analysis and policy evaluations. On one side, external researchers might collaborate in public organisations, as part of the internal staff (through internships or consultancy positions), generating improvement in the quality and accuracy of the research and bringing analytical skills that can potentially create knowledge spill-overs to internal staff. On the other side, researchers can access data to conduct impartial policy evaluations that may contribute to support policymaking “from the outside”.

To facilitate the process of data access, **a common taxonomy for legal, technical and organisational rules** has been recognised internationally, using “annodata” (INEXDA, 2020_[52]; Bender et al., 2019_[53]). Moreover, the use of **common metadata across datasets** and standardised procedures to request data access is crucial to simplify the procedures and give researchers common indications on data sources.

Data can be shared through **public-use files** and **research-use files**. The benefit of public-use files is straightforward, since data can be easily accessed online. Yet, they usually contain less, and non-confidential, information (e.g., anonymised records) than raw data. Research-use file thus allow to access more comprehensive and sensitive information, while preserving confidentiality – using (*scientific-use file*) or not statistical disclosure control (*secure-use files*) – enabling more detailed policy analysis.

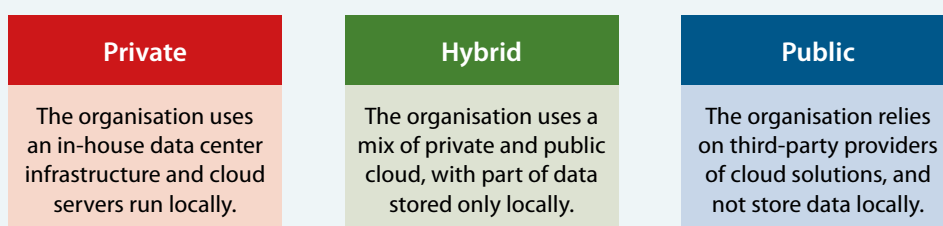
processes (leading to maximum convenience). PaaS provides a more structured platform to deploy own applications and services, and users do not manage or control the underlying infrastructure such as networks or operating systems, with the service provider managing virtualisation operations (see Figure 4.3).

FIGURE 4.3. Categories of Cloud service models and implied characteristics



An organisation using a cloud-based solution may either use an **in-house data centre** infrastructure and run private cloud locally or rely on external providers of cloud servers. The combination of in-house servers and cloud-based solutions (hybrid approach) may enable to meet different organisational needs, both in terms of data security and sharing (Figure 4.4).

FIGURE 4.4. Different cloud deployment models



Source: (IBM, 2019_[50]; Microsoft, 2022_[51]; OECD, 2014_[49])

International framework and practices have defined standards and methods to allow safe micro-data access to research. Several countries are following the **Five Safe Framework**, an internationally recognised guidance to designing micro-data access to research. The Five Safe Framework builds data access around five areas: *safe people*, *safe projects*, *safe settings*, *safe data*, *safe outputs*.⁵ This framework is widely used by national statistical agencies (such as the Australian Bureau of Statistics, Statistics Canada, Stats NZ), research institutions (such as the IAB in Germany), provinces and individual agencies (e.g., Province of British Columbia, BC Ministry of Citizens Services).

Governments and Statistical Offices provide micro-data access for research mainly in three – non mutually exclusive – ways: **on-site access**, **remote access**, and **remote execution**. Pros and cons of these modes are analysed in Table 4.3, including examples of countries – and relative institutions or Research Data Centres (RDCs) – that use them. Remote access seems the most convenient solution to ensure secure micro-data access for scientific purposes, even if privacy regulation is still a major obstacle to the use of remote access for some countries (OECD, 2014_[54]) as it allows to preserve confidentiality while enabling researchers to work from their own device – without being on-site.

A common practice for data access is to grant access on a project-basis: permission will only cover specific datasets, for a defined project or research. However, this limits the researcher’s original scope of the analysis, and precludes linkages across research areas.

5. *Safe people* (it defines that researchers need to be trusted to obtain access to confidential data), *safe projects* (it defines the way in which data provider need to assesses projects’ appropriateness), *safe settings* (it defines how to guarantee a secure environment for data access), *safe data* (it defines how to protect data, e.g. through de-identified information), *safe outputs* (it defines how to ensure that researchers do not disclose confidential information when publishing their research).

Several institutions in Europe have been successful in supplying microdata and providing access for research (see Table 4.3). Among them, a case study for the **RDC of the IAB in Germany** is discussed in Box 4.7.

TABLE 4.3. Mode of micro-data access for research

Mode of access	Pro	Cons	Some improvements	Countries ⁶
On-site access	<p>Researchers can directly see the data</p> <p>High degree of data security: data can only be accessed through local devices and data cannot be download</p>	<p>High cost for researchers: need to travel</p> <p>Only few slots are typically available and need to be booked in advance</p>	<p>Creation of a network of on-site access centers in local offices or universities to increase the number of access locations</p>	<p>First mode of access and the most widely used even now. Examples : INE – Portugal CB; Eurostat's centre in Luxembourg; IAB Germany; Statistics Netherlands</p>
Remote execution	<p>High level of security: researchers do not access the data. They draft the codes that are executed by the data provider</p> <p>No need to travel</p>	<p>Researchers cannot see the data but only final outputs</p> <p>Not suited for complex statistical processing</p> <p>High cost of data preparation – data cleaning and structuring – for the data provider</p>	<p>Provide researchers with a synthetic dataset with similar characteristics but no actual records</p> <p>Allow one-time on-site access before using remote execution system</p>	<p>LIS Luxembourg (Lissy System); StatCan (RTRA system); IAB Germany (JoSua software); Australian Bureau of Statistics (RADL system); Statistics Netherlands</p>
Remote access	<p>Researchers see the data and access them from their own computers</p> <p>No need to travel</p> <p>Safe access using VPN connections</p>	<p>Privacy regulations limit some country's ability to implement this system</p> <p>Relative to remote execution there is a higher risk of disclosure sensitive data</p>	<p>Apply restrictions on some variables and software commands to reduce the privacy concern</p>	<p>Statistics Denmark; Statistics Canada; CBS Netherlands; CASD France; UKDS; Statistics Finland (FIONA system); IAB Germany</p>

Source: (Silberman and Alkhoury, 2021_[55]), (Cole et al., 2020_[56]), (Crato and Paruolo, 2019_[36]), (Desai, 2014_[57])

One of the most renowned national best practices for data access to research is provided by the **VisitINPS** scholar project. As described in Box 4.8, this programme enables national and international scholars to access the entire INPS archives⁷ during visiting periods at INPS's headquarters. Scholars can present a research project on one of the broad topics identified by the INPS Scientific committee (or on another subject related to the Institute's "Strategic Research Areas") and leverage the microdata available in INPS archives for research purposes.

In Estonia, before 2010, **Statistics Estonia** did not provide regulated and systematised excess to microdata. In 2010, the Ministry of Finance, the Ministry of Education and Research, the Bank of Estonia and Statistics Estonia organised study visits to other countries Statistical Offices – Luxemburg Income Study, Statistic Finland, Statistic Denmark and Statistics Netherlands – to implement a new policy and related procedures for disseminating microdata for research. Statistics Estonia successfully implemented such best practices and adapted them to national characteristics, and now has one of the world's most liberal policies for the access of data (OECD, 2014_[54]).

Over recent years, several countries have also collaborated internationally to **share experiences in granting access for research**, creating international networks such as INEXDA Network among Central Banks, Eurostat and international organisations. Similarly, collaborations have been established to increase data sharing across countries, e.g., the "Nordic Microdata Access Network" (NMAN),⁸ the "International Data Access Network" (IDAN),⁹ and the Nordic Institute for Interoperability Solutions (NIIS).

6. The list of countries should be considered as some examples and not an exhaustive list.

7. INPS has the richest data archives on the Italian social security system, pensions, labour market, income support measures and, more generally, welfare-related issues. <https://www.inps.it/dati-ricerche-e-bilanci/attivita-di-ricerca/programma-visitinps-scholars>.

8. The NMAN project, ended in 2017, was built with the aim of giving researchers access to individual-level data from several Nordic countries: Denmark, Finland, Norway, Sweden, Greenland and Iceland (UNECE conference 2021).

9. The IDAN is the collaboration of six RDCs: CASD (in France), IAB and GESIS (Germany), CBS (Netherlands), ONS and UKDS (UK), which build reciprocal data access points in these countries. <https://www.casd.eu/en/le-centre-dacces-securise-aux-donnees-casd/reseau-europeen-idan/>

BOX 4.7: THE RESEARCH DATA CENTRE AT THE INSTITUTE FOR EMPLOYMENT RESEARCH (RDC-IAB) OF THE FEDERAL EMPLOYMENT AGENCY – GERMANY

The RDC-IAB is one of the most successful institutions in Europe in providing access to microdata for research. The RDC is established at the Federal Employment Agency (BA) within the Institute for Employment Research (IAB) in Germany.

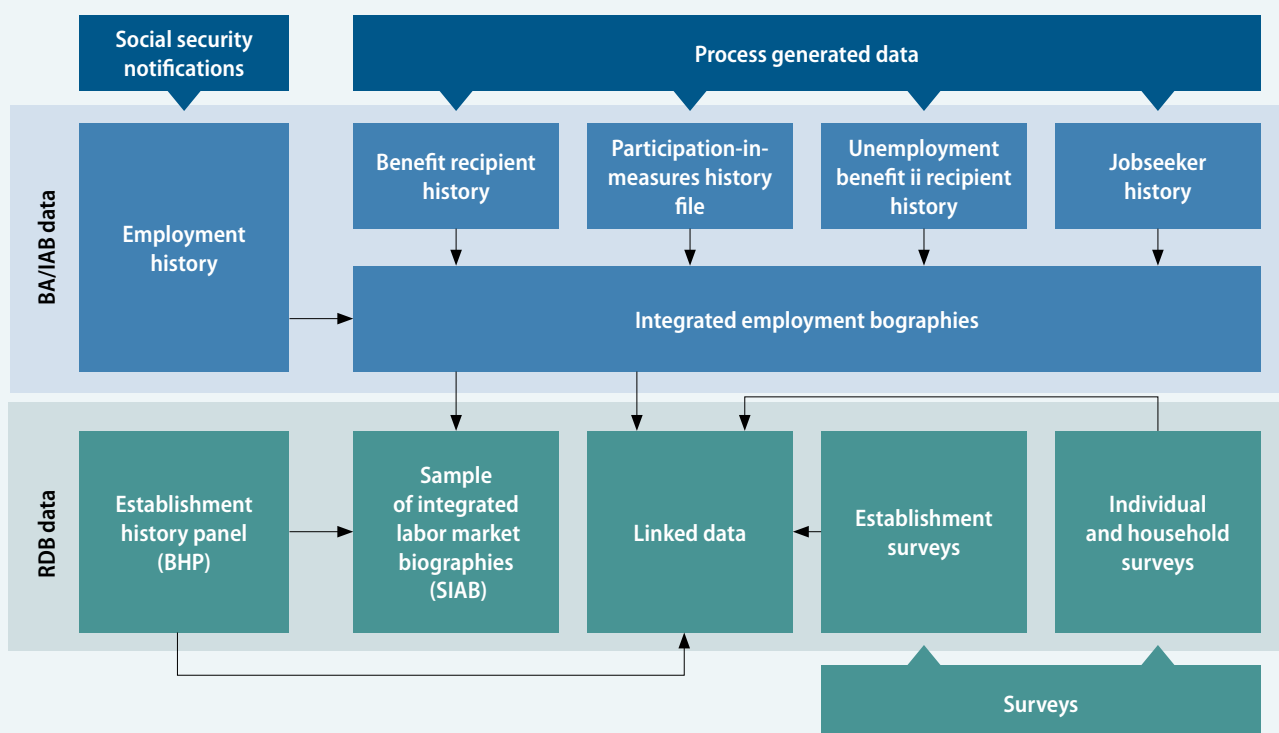
The RDC-IAB provides researchers with the access to data in the fields of social security and labour market, including **individual, household, and establishment data**. Data are collected from administrative sources (social security data and data from internal procedures of the BA) and surveys conducted by IAB or partner institutions. The RDC-IAB integrates data on individuals (in the Integrated Employment Biographies) and establishments (IAB Establishment Panel) and link them to obtain **linked employer-employee data** that allow detailed analysis on both demand and supply side of the German labour market. Linked data are obtained integrating both administrative and survey data (see Figure 4.5).

Access to the data is granted in different ways, responding to different degree of information accessibility: anonymised data (*Campus* and *Scientific-use file*); de-identified data (*Weakly anonymised data*), that offers more sensitive information via on-site access – in Germany and abroad; *remote execution* using the Job Submission Application (JoSua). JoSua is an innovative software that provides separate modules for job submission, distinguishing between internal use (IU) mode – where researchers can upload their code and preview the results – and a presentation/publication (PP) mode – used when data preparation is done, and all outputs and files will be checked by the RDC-IAB staff. Remote execution is possible also after on-site access.

The RDC-IAB provides detailed metadata for all data provided, both in German and in English. The RDC-IAB is thus structured to provide data access also to international researchers, including physical locations in the USA, Canada and UK.

RDC-IAB data have been used for **labour market research and evaluations of specific policies**. In several cases, research conducted using RDC-IAB data have provided **evidence-based advice** for the German government. One example is the analysis conducted using RDC-IAB data on the effects of the minimum wage introduced in 2015: the results was used by the Minimum Wage Commission in their decision to adapt the minimum wage.

FIGURE 4.5. BA/IAB data sources



Source: (Müller and Möller, 2019_[58]), (Müller and vom Berge, 2020_[59]), RDC-IAB website: <https://fdz.iab.de/en/startseite-en/>

BOX 4.8: VISITING SCHOLARS AT INPS: A WAY PROGRAMME TO ATTRACT RESEARCHERS

In 2015, INPS launched its “VisitINPS scholars” programme, through which selected researchers are able to spend a period of study at INPS, with the aim of promoting the use of INPS administrative archives (e.g. Italy's richest data archives on the social security system, pensions, labour market, income support measures and more generally on welfare related issues) and providing analysis and evaluation in strategic research areas.

To participate to this project, researchers are asked to submit a research project on one of the broad topics of interest to INPS, which is then assessed by a Scientific Committee. Once the project is completed, it is published inside INPS as a working paper and the researcher has the right to publish it in academic journals.

The VisitINPS Scholars programme benefits from the financial support of a number of Universities, Financial Institutions, Foundations, private companies and research institutes.

Based on the positive experience of the VisitINPS programme, in 2018, the INPS management created its own dedicated Research Department (Direzione Centrale Studi e Ricerche), which has been growing gradually and now embeds four professors as full-time INPS staff. Their role is to directly support INPS president and board through evidence-based advice, using research to improve INPS policies and administrative procedures.

Source: (INPS, 2021^[60])

Building strong relationships with academia and research centres is crucial. An example on successful partnerships between state government and the research community can be appreciated from the Ohio Longitudinal Data Archive (**OLDA**) – that will be further discussed in Box 4.10 – and from the **New Brunswick Institute for Research, Data and Training (NB-IRDT)** in Canada (discussed in Box 4.9). Both cases provide good examples on how leveraging skills and infrastructure from academia can be an effective solution when data providers lack internal capacity to maintain the data access infrastructure and to conduct policy analysis.

BOX 4.9: NEW BRUNSWICK INSTITUTE FOR RESEARCH, DATA AND TRAINING (NB-IRDT) – CANADA

The NB-IRDT is a provincial Research Data Centre and data custodian resulting from a long-term collaboration between academia and government. The collaboration has included several government departments such as the Dep. of Health, the Executive Council Office, the Dep. of Post-secondary Education, Training and Labour and the Dep. of Social Development.

The NB-IRDT collects data from government agencies and non-profit organisations, and it provides research access to over 45 linkable personnel and health datasets (including microdata on health, labour, training and education) to conduct research that support evidence-based policy.

While the NB-IRDT is situated in the University of New Brunswick, the government keeps an active engagement in setting research priorities and collaborative programs that focus on crucial policy areas.

Source: <https://www.unb.ca/nbirdt/>; (Maillet and Ted McDonald, 2020^[61])

DEVELOPING EFFECTIVE DATA VISUALISATION TOOLS TO SUPPORT EVIDENCE-BASED POLICYMAKING

Data visualisation is a key element of evidence-based policy making: it allows to effectively communicate evidence to policymakers – and to the public at large – in a comprehensive, timely and accessible way. Data visualisation increases transparency on the use of public resources and on the effectiveness of public policies (Umbach, Guidi and Russo, 2018^[62]).

Among international best practices on the use of data visualisation tools to support policymaking, the **Workforce Success Measures dashboard (WSMD)** displays outcomes for Ohio's largest workforce programs in the US. The dashboard is managed by the Ohio State University: it is, thus, a good example on how collaboration between government agencies and academia can lead to the creation of a powerful tool to compare the effectiveness of different programs and support decision making (see Box 4.10 for more details).

The case of WSMD proves the importance of providing relevant information and consistent indicators across time and policy measures. To guarantee that visualisation tools are accessible and understandable, it is crucial that figures and tables are complemented with detailed explanation of indicators and data used. ■

BOX 4.10: WORKFORCE SUCCESS MEASURES: A DASHBOARD OF WORKFORCE PROGRAM PARTICIPATION TO SUPPORT POLICY-MAKERS – OHIO, US

The dashboard is managed by the Ohio Education Research Centre (OERC) of Ohio University, which has worked in collaboration with the Governor's Office for Workforce Transformation (OWT). It displays workforce training outcomes over time for the state's largest workforce programs in Ohio.

The dashboard uses common outcomes/indicators to measure different training and education programs, with the aim of giving policymakers coherent tools to compare outcomes across programs and evaluate their effectiveness. For each program, the dashboard shows: the number of individuals that benefits from the intervention, how many have subsequently found a job in Ohio, their median earnings, their employment stability before/after completing the program, and information on education.

The outcomes showed in the dashboard are calculated using administrative data from the Ohio Longitudinal Data Archive (OLDA), a centralised database that contains longitudinal data from several State agencies as well as data from the private sector. On several occasions, the State has commissioned to the University the production of specific studies on States' policies using OLDA data.

Source: (Hawley, 2020^[63]); <https://workforcesuccess.chrr.ohio-state.edu/home>

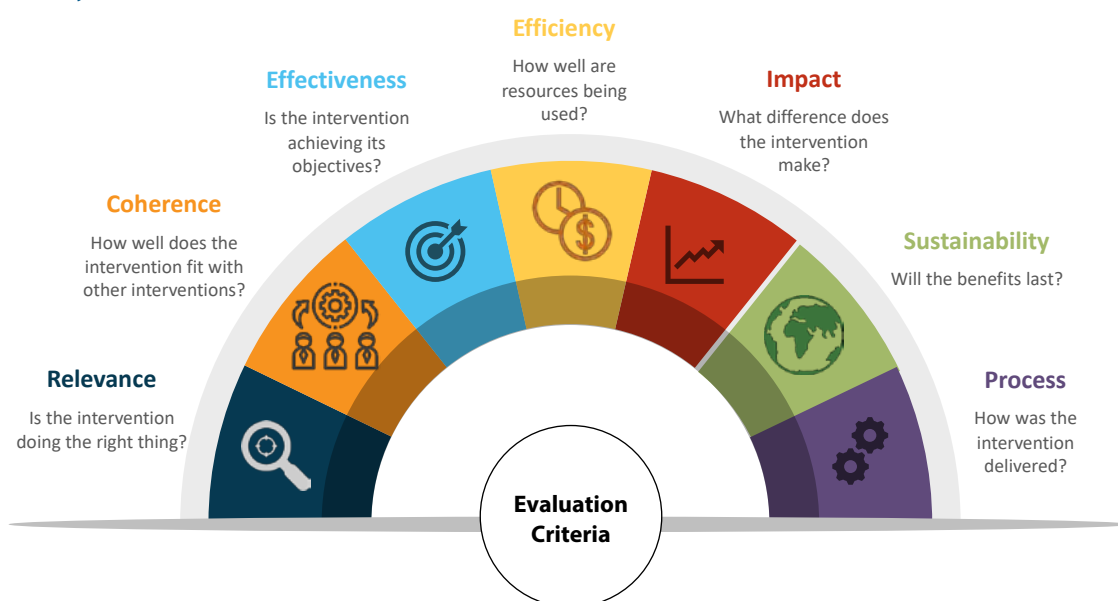
5. Developing the practice of quantifying policy impacts: key methodological considerations and international best practices

This section provides some considerations on how to approach impact evaluation and the quantification of policy impacts. When evaluating policies, it is particularly important to plan early what should be evaluated, what data and evidence should be collected before implementation and during the timeframe of the intervention. If these activities are left until after the lifetime of the intervention, this may limit the ability to apply appropriate methodologies and obtain meaningful results from an evaluation. The section looks first at some of the questions that policy evaluations can help answer, as well as the methodologies that can be used to find answers. Finally, the section delves into actual applications of these methods are presented with a series of international examples.

IDENTIFYING THE KEY QUESTIONS TO ANSWER IN THE IMPACT EVALUATION

The impact of an intervention can be assessed from multiple angles, each calling for distinct methodologies for evaluation. In line with the literature and best practice on impact evaluations (Box 5.2 below), seven broad criteria can be identified to evaluate an intervention (Figure 5.1), which correspond to different sets of evaluation questions.

FIGURE 5.1. Key evaluation criteria



Source: OECD elaboration, integrating the OECD DAC Criteria (OECD, 2021^[64]), OECD work on principles and standards for policy design, implementation and evaluation (OECD, 2020^[65]) and the discussion on Process evaluation in the UK Magenta Book (HM Treasury, 2007^[66])

The criteria should be viewed as a set of lenses through which one can understand and analyse an intervention. They provide a classification of the main evaluation questions that policy makers usually consider. Together they highlight how interventions should be relevant to the context, coherent with other interventions, achieve results in an efficient process and have positive, lasting impacts for sustainable development. The criteria are related and can be seen as a complementary set. The definition of each criterion presents a distinct concept and yet these concepts are in many ways interrelated.

Table 5.1 provides a summary of the evaluation criteria and the respective evaluation questions. The evaluation questions will slightly differ in accordance with whether the evaluation is conducted ex-ante, in-itinere or ex-post.

TABLE 5.1. Overview of evaluation criteria and sample evaluation questions

Evaluation Criteria	Concept	Sample Questions
Relevance: Is the intervention doing the right things?	Relevance refers to the extent to which the intervention's objectives and design respond to beneficiaries' global, country and partner/institution needs, policies and priorities, and continue to do so if circumstances change.	<ul style="list-style-type: none"> To what extent are the intervention's objectives in line with the needs of the country, region, and society? Is the intervention in line with the needs of the target group? Is the intervention adequate as a strategy to produce an effect with respect to the issues of the target field and sector in the country?
Coherence: How well does the intervention fit with other interventions?	Coherence refers to the compatibility of the intervention with other interventions in a country, sector, or institution. This is assessed both in terms of <i>internal coherence</i> (coherence with other interventions by the same institution) as well as <i>external coherence</i> (i.e., consistency with other actors' interventions in the same context).	<ul style="list-style-type: none"> Is the intervention compatible with other pre-existing or planned interventions in the country, sector and institution, including ones that have been introduced after the launch of the intervention? Which other interventions support or undermine the intervention, and vice versa?
Effectiveness: Is the intervention achieving its objectives?	Effectiveness refers to the extent to which the intervention achieved, or is expected to achieve, its objectives and its results.	<ul style="list-style-type: none"> Is the project objective clearly mentioned? Do the indicators for the project objective accurately express the objective? To what extent have the interventions' objectives and milestones been achieved? Have the interventions and instruments used produced the expected effects? Could more effects be obtained by using different instruments? Are there factors that impede the achievement of the objectives?
Efficiency: How well are resources being used?	Efficiency refers to the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.	<ul style="list-style-type: none"> Is this intervention a good use of resources? Have the planned outputs been achieved at the lowest cost? Are there alternative uses of resources that would yield more benefit? What costs and benefits does the intervention generate for different groups in the population?
Impact: What difference does the intervention make?	Impact refers to the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.	<ul style="list-style-type: none"> What measurable outcomes, both intended and unintended, occurred? How much of these outcomes can be attributed to the intervention? Have different groups been impacted in different ways, how and why? How has the context influenced outcomes? Can the intervention be reproduced?
Sustainability: Will the benefits last?	The sustainability criteria look at the extent to which the net benefits of the intervention continue or are likely to continue.	<ul style="list-style-type: none"> To what extent will the net benefits of the intervention continue after the intervention's end? Will they continue if there is no more public funding? What is the evidence on the sustainability of the intervention? What are the financial, economic, social, environmental and institutional capacities of the systems needed to sustain net benefits over time? What are the long-term risks and trade-offs of the intervention? Can the intervention be reproduced? What generalisable lessons have we learned about impact?
Process: What can be learned from how the intervention was delivered?	The process criterion aims to analyse the different steps of the policy implementation and identify any dysfunctional decision points that inhibited optimal rollout and delivery.	<ul style="list-style-type: none"> What worked well and less well, and why? What could be improved? How has the context influenced delivery?

SELECTING THE APPROPRIATE METHODS

A wide range of evaluation methods exists to answer different evaluation questions and quantify the impact of a public policy. Different methods present different trade-offs and implications on resources, data requirements and timing of the evaluation.

Impact evaluations usually rely on a mix of methods, both qualitative and quantitative

The selection and implementation of different methods will depend on the assumptions on how the policy intervention will produce a desired outcome on beneficiaries. It will also hinge crucially on data availability. There should be alignment between the evaluation questions, the evaluation approach and the chosen quantification methods. In practice, a large number of impact evaluations rely on a mix of methods, combining both qualitative and quantitative approaches to answer questions about whether a policy intervention was implemented correctly, whether it produced the desired outcome and whether it was cost-effective. No individual method can provide answers to all evaluation questions. If time and resources are limited, questions and methods will have to be prioritised and trade-offs between methods will need to be assessed.

It is important to consider, the feasibility, appropriateness and timing of evaluations

The cost, timing, respondent burden, ethics, response rate, and the potential effect of data collection itself should all be considered (HM Treasury, 2020_[67]). Clear definitions should be developed from the start around the targets, places and time periods associated with the evaluation questions to ensure that the most appropriate methods are identified. It is also important to have a spatial dimension of an evaluation, if relevant. This dimension may be stronger in some programmes than others (e.g., regional programmes), and as such should be clearly identified ahead of any impact assessment (EC, 2013_[2]).

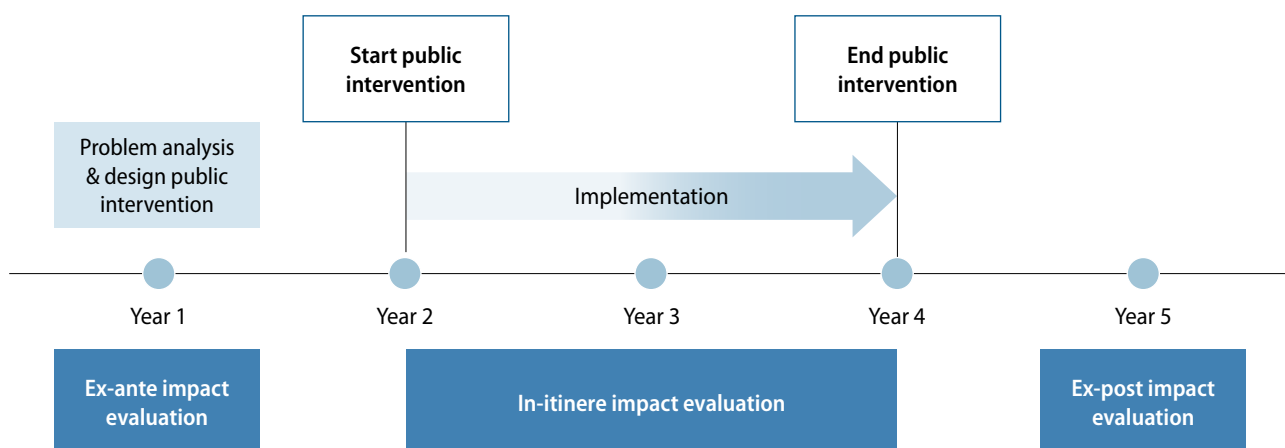
Moreover, when designing an impact evaluation, it is particularly important to consider the timing of the data collection, to align it with the key implementation milestones. Evaluations' objectives, data, methods and usage, vary depending on the point of the policy cycle at which evaluations are carried out. Three broad categories can be identified:

- **Ex ante:** before an intervention is fully formed, evaluation should be used to help shape its design and how it will be implemented. Methods to be considered include analysing evaluation evidence from other similar interventions, conducting simulations, piloting and early testing of policy ideas. An early testing would include an experimental design, indicating the potential beneficiaries detected in the baseline, who will receive treatment in the following years. The evaluation can help in identifying expectations about the intervention's impact; possible implementation bottlenecks and risks and baseline evidence that can be used to evaluate the intervention once implemented.
- **In itinere:** evaluating an intervention under implementation offers the greatest opportunity to influence decisions and help ensure that the policy can realise its intended benefits. These evaluations will typically collect evidence about the efficacy of the policy's design, its implementation and emerging outcomes. They can examine questions such as policy delivery and impact, differential impact of the policy on different target groups, early indications of possible effect size, monitoring unintended consequences and possible 'in-flight adjustments' to the policy.
- **Ex post:** these evaluations can make more conclusive statements on the design, implementation and effect size of the intervention, with information on the beneficiaries of the policy and the outcome variable that the policy intended to improve. Many of the techniques used for ex-ante and in-itinere evaluations can also be used for ex-post evaluations. However, there would be instances where the conditions to conduct an experimental design are not met and quasi-experimental designs will be adopted (e.g., matching, difference-in-difference, instrumental variables, regression discontinuity methods).

Choosing the right timing of the evaluation will also impact the quantification methods that can be adopted and is critical to ensure that the evaluation delivers useful, high-quality findings to those designing and implementing policies and the right data are collected in the most cost-effective and efficient way possible (Figure 5.2). It is important

to note, however, that the distinction between ex ante, in itinere and ex post is not always obvious. For example, there are policy changes that will apply from a certain point onwards and will never be discontinued (e.g., an increase in the unemployment benefit replacement rate). For such policy changes, there is no real “end public intervention” and thus no clear-cut distinction between in-itinere and ex-post.

FIGURE 5.2. Impact evaluation along the phases of policy design and implementation



Source: OECD elaborations.

OVERVIEW OF THE MAIN QUANTITATIVE AND QUALITATIVE EVALUATION METHODS

A simple but intuitive way of classifying evaluation methods is to think of them in terms of the results they can produce, namely quantitative and qualitative. Their use provides different evidence that can be part of an evaluation design. While the distinction made between quantitative and qualitative methods appears as common-sense, it is not always clear-cut. For example, the foundations of quantitative evaluations can sometimes be qualitative (e.g., ‘50% of those interviewed said they had benefited from the programme’). Some qualitative methods, such as systematic reviews, are meant to qualitatively synthesise evidence from quantitative studies. Despite this caveat, the distinction between the two categories remains a useful one for the purpose of classification.

Quantitative approaches

These approaches produce sizes and orders of magnitude of the actual and/or expected effects of a policy intervention. The main alternatives are:

- **Descriptive Statistics:** analyse a large number of cases – e.g., individuals, firms – and their characteristics. This analysis can study the situation or evolution of policy beneficiaries.
 - *When to use them?* By nature, their use is limited to in-itinere and ex-post evaluation.
- **Counterfactual:** this approach compares different situations in which an intervention is present or not, or different intervention options. It estimates what would have happened in the absence of an intervention and compares it with what has been observed thanks to the intervention (Randomised Controlled Trials, difference-in-difference, matching, instrumental variables, regression discontinuity). A key issue in designing, or recreating, such an evaluation is the identification of treatment and control groups (Box 5.1).
 - *When to use them?* Counterfactual experimental designs are predominantly applied for ex-ante and in-itinere evaluations (in the form of pilot RCTs). Quasi-experimental designs that reproduce counterfactual conditions can be applied both in-itinere and ex-post.

- **Simulation models:** Simulations are used to study how the behaviour of different economic actors would change under different policy scenarios. With computer-aided simulations, increasingly complex models can be simulated by designing and developing mathematical representations of complex scenarios. In policy settings where a treatment and control group cannot be created (for example, when a policy has a general application), the counterfactual can be created using mathematical models. This is done by developing a model which simulates a set of outcomes from a set of inputs, by combining data, evidence-based assumptions, and a chosen simulation logic (HM Treasury, 2007_[66]).
 - *When to use them?* While extensively used ex-ante, simulations can also be conducted in-itinere and ex-post.
- **Cost-benefit and cost-effectiveness analysis.** These quantitative methods are used to assess efficiency considerations. Cost Benefit Analysis (CBA) is a method of comparing the relevant costs and benefits of policy decisions. The ultimate objective is to compare costs and benefits to assess whether the intervention represents value-for-money for society. Cost Effectiveness Analysis (CEA) is a variant of CBA which enables policy makers to compare various options in terms of the implied cost of the option, and the likely outcomes from pursuing that option (HM Treasury, 2018_[68]).
 - *When to use them?* Can be applied at any point of the policy cycle.

Qualitative approaches

For these approaches the information used is mainly qualitative, based on observation, documentary analysis and surveys. The main methodologies that can be used in impact assessment are:

- **Theory-based approaches:** these are methods in which the current situation is compared to pre-existing theories or other causal pathways identified during an assessment. Within this approach, the most used options are:
 1. **realist evaluation** articulates specific, hypothesised causal mechanisms, and evidence is gathered for each. This method focuses on not only identifying what works but also how, why and over what duration, i.e., *what works in which contexts* (Westhorp, 2014_[69]).
 2. **process tracing** is a structured method to develop and assess theories on how a particular outcome originated. It analyses a single case of change and tests whether a hypothesised causal mechanism can explain the results (HM Treasury, 2007_[66]).
 3. **contribution analysis.** Contribution Analysis refers to methods which are used to understand the likelihood the intervention has contributed to an outcome observed. It functions through a step-by-step process which analyses how the contribution would have arisen and uses a broad range of evidence to test this (HM Treasury, 2007_[66]).
 - *When to use them?* While realist evaluation can be applied at all phases of an intervention, contribution analysis and process tracing can only be applied in-itinere and ex-post as these methods examine elements from the actual implementation of an intervention.
- **Participatory Evaluations** consider the importance of the different stakeholders (beneficiaries, managers, experts, others) in the intervention. Their opinions, assessments, judgments and experience allow the evaluators to identify the most relevant theories of change. These techniques, like rapid impact evaluations (RIEs), are usually based on interviews/focus groups or surveys.
 - *When to use them?* They can be applied at all stages of an intervention.
- **Synthesis Methods** involve combining and summarising a set of different documents and evaluations to arrive at a judgment based on cumulative findings (systematic reviews and meta-analysis). For example, one common synthesis method is the qualitative comparison of different case studies.
 - *When to use them?* They can be applied at all stages of an intervention.

Each method can respond to different evaluation questions. Table 5.2 below maps the more suitable methodology to the evaluation questions presented above. Overall, this approach offers a novel framework to categorise policy evaluation methods (Box 5.2).

BOX 5.1: THE MAIN CHALLENGES OF COUNTERFACTUAL EVALUATION AND WAYS OF ADDRESSING THEM

Identifying “appropriate” treatment and control groups

This box presents a few key considerations which are needed to assess whether a public intervention – called “D” – has an impact (α) on a certain characteristic of individuals or units like firms, an outcome variable named “Y”.

Once the public intervention has been implemented, two population groups can be distinguished:

- Those individuals/units who have actually received the public intervention, also called beneficiaries, (the treatment group).
- Those individuals/units or organisations that have not received the public intervention (the control group).

An impact evaluation aims to understand whether the intervention made a difference on beneficiaries and to isolate the effect that is due exclusively to the intervention. Ideally, the outcome Y_1 observed among the individuals/units that receive the intervention ($D = 1$) would be compared to the outcome observed if the *same* individuals/units or organisations would have not benefited from the intervention, Y_0 . Comparing the same individuals would ensure that any difference is due to the intervention. In mathematical terms, this approach can be expressed as:

$$\alpha = E(Y_1|D = 1) - E(Y_0|D = 1) \quad (1)$$

However, it is impossible for an evaluator to observe what would have happened to a treated individual in the absence of treatment, Y_0 . The solution is to calculate the average of the outcome variable in the treated group, $E(Y_1|D = 1)$, minus the average result obtained from individuals who have not benefited from the intervention – the control group, $E(Y_0|D = 0)$.

$$dif_{treat-cont} \alpha = E(Y_1|D = 1) - E(Y_0|D = 0) \quad (2)$$

However, these two quantities could be different due to differences in the two groups which do not depend on receiving treatment. The difference between the true impact and the estimate is the so-called selection bias.

$$dif_{treat-cont} = \alpha + bias.selection \quad (3)$$

Therefore, the estimate of the impact of an intervention calculated by the evaluator is equal to the true impact of the intervention, α , only if the selection bias is zero, ($bias.selection = 0$).

Selection bias indicates that individuals in the treatment and control groups are not directly comparable, because there are differences between them (and not just because some receive the intervention and others do not). These differences may stem from ex-ante characteristics (age, gender, level of education, preferences and tastes, etc.), or from contemporaneous factors that affect both the likelihood that the individual receive the intervention and the outcome of interest.

The bias can be addressed through the following strategies:

- **The evaluation design guarantees that the selection bias is zero.** This is the case for experimental designs like Randomised Control Trials (RTC), where the assignment of individuals to the treatment and control group is carried out through a random process, like tossing a coin.
- **The selection bias can be corrected through “quasi-experimental designs”.** A bias correction is achieved by applying econometric and statistical techniques. The most common selection bias correction methods are:
 - **Matching methods** (Propensity Score Matching – PSM), where participants in the treatment group are paired to participants in the control group based on the similarity of their scores to account for selection bias;
 - **Differences in differences** (diff-in-diff), where the effect of a treatment or of a policy is estimated by comparing the pre- and post-treatment differences in the outcome in the treatment and control group;
 - **Instrumental variables method**, where a chosen variable (instrument) which fulfils a set of conditions (i.e. induces changes in the explanatory variable but has no independent effect on the dependent variable) allows evaluators to uncover the causal effect of the explanatory variable on the dependent variable;
 - **Regression Discontinuity**, where participants are assigned to treatment and control groups based on a cut point of an assignment variable. The discontinuity in the output variable between the treatment and control groups at the cut point allows to identify the effect of the policy intervention. The assignment variable (and, specifically, its cut point) can be thought of as an instrumental variable, in that it induces changes in the treatment but does not independently affect the outcome.

BOX 5.2: A NOVEL FRAMEWORK TO CATEGORISE POLICY EVALUATIONS METHODS

Building on the relevant literature and international best practice, this report aims to present a comprehensive mapping which connects evaluation questions (as defined by the seven key criteria), the timing of policy evaluation (ex-ante, in-itinere, and ex-post) and the different quantitative and qualitative methodologies to address these.

In particular, the seven criteria, build on the recognition that many sources converge on the key evaluation criteria. For example:

- EVALSED uses: 1) Relevance, 2) effectiveness, 3) efficiency, 4) utility and sustainability
- The Italian Guida all'analisi e alla verifica dell'impatto della regolamentazione uses: 1) Relevance, 2) coherence, 3) efficiency 4) effectiveness
- The Nordic Council's Policy Instrument on evaluation mentions: Relevance, efficiency, effectiveness
- JPCA: Relevance, Effectiveness, Efficiency, Impact, Sustainability
- The Magenta Book: 1) Process, 2) impact, 3) efficiency (i.e., value for money)
- Japan Ministry of Foreign affairs: relevance, effectiveness, appropriateness of process
- OECD work on principles and standards for policy design, implementation and evaluation (OECD, 2020_[65]) informing the 2022 OECD Council Recommendation on public policy evaluation

The framework here presented reflects these common criteria and it builds on the six comprehensive OECD DAC criteria (Relevance, Coherence, Effectiveness, Efficiency, Impact, Sustainability) to complement them with one additional aspect (i.e. Process) which appears in the Magenta Book. In addition, the framework distinguishes evaluation questions between ex-ante, ex-post and the often neglected in-itinere timings, while also keeping the distinction between qualitative and quantitative methodology. The framework is meant as a single tool that MIMIT can help to decide on appropriate evaluation methods for its policies.

QUANTIFYING POLICY IMPACTS: EXAMPLES FROM OECD COUNTRIES

Lessons from the experiences of policy applications are especially useful to anticipate potential challenges. This section presents case studies of ex-ante, in-itinere and ex-post policy evaluations in OECD countries, focusing on selected methodologies. The case studies spell out the policy issue at stake, the methodological challenges and how these challenges were addressed. Advantages, limitations and lessons are presented for each method to start building a set of guiding principles that MIMIT's Centro Study can consider when quantifying policy impacts.

Ex-ante impact evaluation

Ex-ante impact evaluations have the primary objective of clarifying and verifying the logical reasoning that links the problem, its underlying causes, the objectives and a range of policy instruments to tackle the problem (Nordic Council of Ministers, 2019_[70]). They must present the likely impacts of the policy instruments, who will be affected by them and how. Ex-ante evaluations must compare the policy instruments on the basis of their assessed economic, social and environmental impacts.

In this section we present case studies from four of the main methodologies which can be applied ex-ante: counterfactual experimental designs, simulations, cost-benefit analysis and systematic reviews.

Counterfactual experimental designs

As introduced in the previous section, the key feature of a Randomised Controlled Trial (RCT) is the use of a random assignment to create at least two groups that closely resemble each other. The only difference is that one group is exposed to a new policy intervention, while the other does not (Figure 5.3). By assessing the differences between the

TABLE 5.2. Overview of evaluation criteria and methods

Criteria	Quantitative Methods	Qualitative Methods
Relevance: Is the intervention doing the right thing?	<ul style="list-style-type: none"> • Simulations • Descriptive statistics (output/performance monitoring) 	<ul style="list-style-type: none"> • Qualitative assessments • Participatory evaluation, e.g., Interviews and focus groups; Surveys and polling • Theory based methods; e.g. Realist evaluation; Contribution analysis; Process tracing
Coherence: How well does the intervention fit with other interventions?	<ul style="list-style-type: none"> • Simulations of interactions between different policies • Quasi-experimental designs 	<ul style="list-style-type: none"> • Qualitative assessments of the legislative environment and existing policies based on the available sources of data/information • Participatory evaluation (Interviews and focus groups; Surveys and polling) • Synthesis methods
Effectiveness: Is the intervention achieving its objectives?	<ul style="list-style-type: none"> • Simulations • Counterfactual experiments: Pilot RCTs • Descriptive statistics • Quasi-experimental designs 	<ul style="list-style-type: none"> • Qualitative assessments • Participatory evaluation (Interviews and focus groups; Surveys and polling) • Synthesis methods • Theory based methods; e.g. Realist evaluation; Contribution analysis; Process tracing
Efficiency: How well are resources being used?	<ul style="list-style-type: none"> • Cost-benefit analysis • Cost-effectiveness analysis • Descriptive statistics 	<ul style="list-style-type: none"> • Synthesis methods (e.g. case studies) • Participatory evaluation (Interviews and focus groups; Surveys and polling)
Impact: What difference does the intervention make?	<ul style="list-style-type: none"> • Simulations • Counterfactual experiments: Pilot RCTs • Descriptive statistics • Quasi-experimental designs 	<ul style="list-style-type: none"> • Qualitative assessments • Participatory evaluation (Interviews and focus groups; Surveys and polling) • Synthesis methods • Theory based methods; e.g. Realist evaluation; Contribution analysis; Process tracing
Sustainability: Will the benefits last?	<ul style="list-style-type: none"> • Simulations • Cost-benefit analysis 	<ul style="list-style-type: none"> • Theory based methods • Synthesis methods (e.g., case studies)
Process: What can we learn from the intervention delivery?	<ul style="list-style-type: none"> • Cost-benefit analysis • Descriptive statistics (Output and performance monitoring) 	<ul style="list-style-type: none"> • Participatory evaluation (Interviews and focus groups; Surveys and polling) • Synthesis methods (e.g., case studies) • Theory based methods; e.g. Process tracing

Note: Text in **black** refers to methods that can be applied ex-ante, in-itinere and ex-post. Text in **blue** to methods that can be applied only ex-ante, and in-itinere. Text in **green** to methods that can be applied only in-itinere and ex-post.

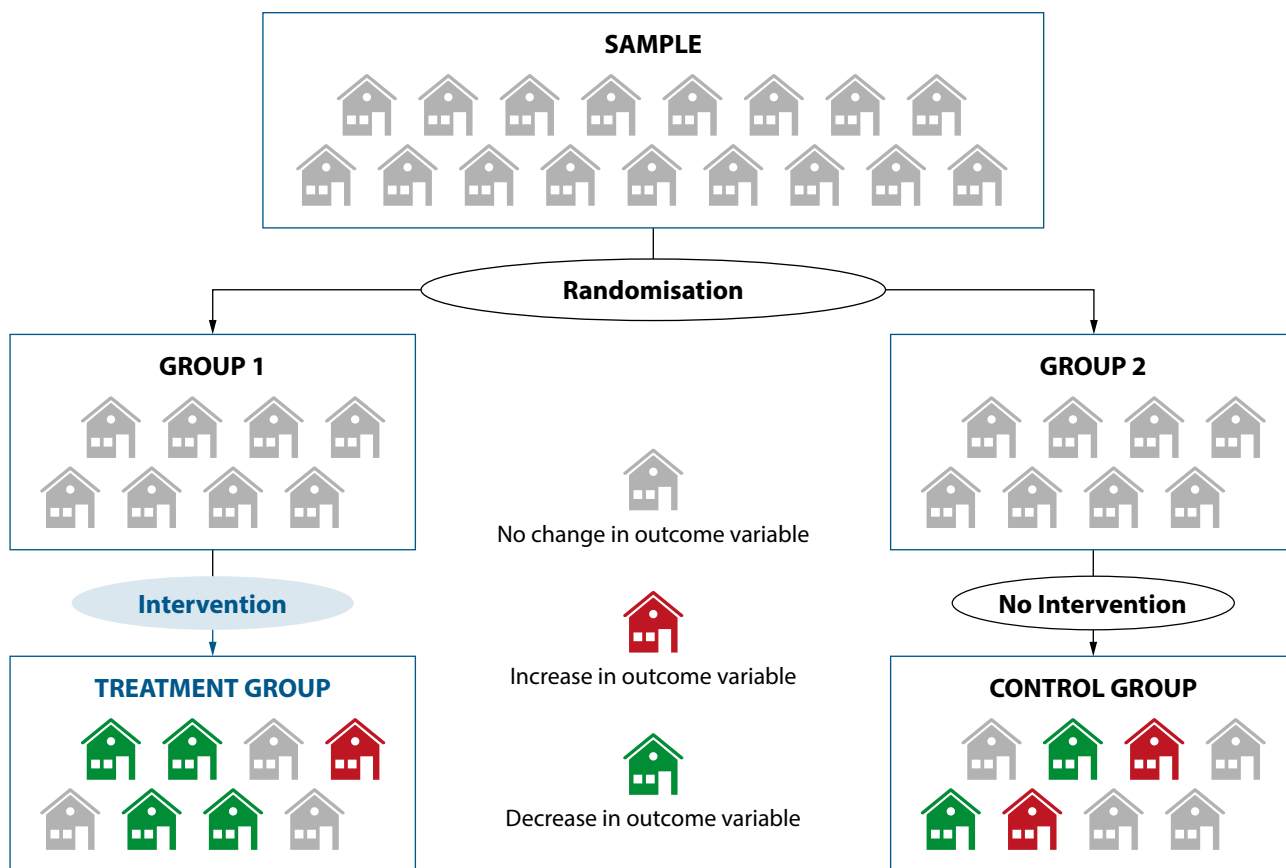
two identical groups, chosen at random, an RCT enables policymakers to understand which strategies, if any, are working, and to eliminate selection biases, pre-existing or external factors that complicate the evaluation process.

The necessary conditions of an “experimental design” are:

- The evaluation is designed at the same time as the public intervention.
- The evaluator randomly assigns individuals to a treatment and control group.
- Once an individual has been selected in a group, no one can change the group.

The second condition (random assignment to treatment and control groups) can be achieved through different methodologies, also depending on the policy design. For example, evaluators can randomly assign half of the population to treatment and half to control (one-stage design); or else they can first randomly draw a representative sample from that population and then randomly assign each individual in the sample to either control or treatment (two-stages design).

FIGURE 5.3. Visualisation of a simple RCT



Source: Elaboration from (OECD, 2019_[71]).

Overall, RCTs are a rigorous and unbiased way to estimate an intervention’s causal impact, as they minimise the risk of confounding factors. However, RCTs may not always be logistically, ethically or politically feasible (OECD, 2019_[71]). Limitations include data requirements (need for baseline and follow-up data for both treatment and control), as well as external and internal validity problems (e.g., due to attrition or failures in random assignments). In addition, RCTs may also involve high costs related to setting up and implementing the experimentation. Online experiments can be used to mitigate costs.

TABLE 5.3. Case study: evaluating business response rates to redesigned institutional letters in Denmark

Policy Issue	The Danish Business Authority faces low response rates when sending letters to businesses, resulting in additional expenses for follow-up letters and reminders. They conducted an experiment to increase the response rate of a specific letter urging businesses to sign up to the public database Nutrition Base. Before the intervention, only half of the businesses that received the letter replied and even less signed up. The experiment tested whether redesigning the letter can increase the response rate, and it was carried out by the Danish Business Authority and the Ministry of Taxation (Danish Business Authority, 2013 _[72]).
Methodology	The experiment involved sending out three different letters to businesses: the original letter (control), a new letter on white paper, and the same new letter on red paper. The new letter used strategies from behavioral economics, including addressing the business by name and emphasizing the cost of not responding. The letters were sent out to three randomly selected groups of businesses, and almost 600 businesses received a letter (200 for each version).
Results	The experiment found that the redesigned letters significantly increased the response rate. Only 57% of businesses responded to the original letter, while 65% responded to the new letter on white paper and 69% responded to the new letter on red paper. Moreover, a higher percentage of businesses that received the new letters actually signed up to the Nutrition Base. The results were used to make a business case for implementing the new letter as the standard one, estimating its costs and benefits.

Source: (Danish Business Authority, 2013_[72]).

Simulation models are quantitative tools that are designed to answer “What if...?” questions when there are different alternatives for implementing an intervention. Simulations are a useful tool for predicting the impacts of policy changes on individuals, including identifying winners and losers, and generating alternative scenarios for policy design. Simulations are particularly useful for examining interactions between policies.

In its simplest form, a simulation is an econometric or mathematical model that describes, through equations, the causal relationship between the variables. In general, they show how a series of factors, characteristics or regressors (D, X1, X2,...), influence an outcome variable (Y) (Table 5.4).

TABLE 5.4. Steps of impact estimation for simulation models

Steps	Description
1	Estimate the econometric model
2	Predict the value of Y in the base scenario
3	Manipulate the value of regressor(s) and estimate Y in reform scenario
4	Calculate impact estimate as difference of the two Y values.

Source: OECD elaboration.

There are many types of simulation models. For example, some adopt a partial approach, where the analysis focuses on an economic agent (e.g., households or individuals), while keeping fixed the rest of the economic agents (companies or the public sector). Such models include: (a) aggregate accounting models that use macro variables, and (b) microsimulation models with information at the disaggregated level of individuals, households or companies. A more general approach is (c) relying on General Equilibrium Models, which simultaneously consider supply and demand in the analyses.

While simulations are relatively inexpensive to run once a model is developed, they are limited by the quality of data and assumptions used in the model, as well as the time and expense required to construct accurate models and, when needed, develop specialised softwares (HM Treasury, 2007_[66]).

TABLE 5.5. Case study: Ex-Ante Evaluation of Investments in Knowledge, Learning and Innovation in Germany

Policy Issue	Knowledge, learning, and innovation are crucial for the success of companies and economies, but it's not fully understood how companies can efficiently exchange knowledge. This German case study (Müller, Kudic and Pyka, 2017 _[73]) shows how agent-based modelling and simulation (ABMS) can be used to evaluate innovation policy interventions in regional economic systems. Simulation allows examining the effectiveness of different investment alternatives in knowledge, learning, and education under realistic conditions before implementing them.
Methodology	The study used agent-based modeling and simulation (ABMS) to evaluate potential innovation policies in a network of 558 interconnected firms in a regional innovation system. The simulation compared a control scenario with no intervention to three scenarios in which 100 new links were generated between small, medium, or large firms. The simulation analyzed the impact of these scenarios on the exchange of knowledge between firms. The study used primary and secondary data collections, and the location of firms was analyzed to assess specific regional effects. Specifically, the simulation looked at what would happen i) in the scenario in which new links are generated among small firms, ii) the scenario in which new links are created among medium firms; iii) the scenario in which new links are created among large firms.
Results	The simulation found that creating new links among small firms was the most effective intervention, and that the network structure played a major role in knowledge diffusion processes. The study demonstrates that even simple models can provide insights for ex-ante evaluation of innovation policy interventions in regional economies. It also emphasizes the importance of a differentiated funding strategy that takes into account region-specific system characteristics.

Source: (Müller, Kudic and Pyka, 2017_[73])

Cost-Benefit Analysis

Cost-Benefit Analysis (CBA) is a method of comparing the relevant costs and benefits of policy interventions to assess whether they represent good value-for-money. The costs and benefits include the financial, environmental, and social consequences accruing to society as a result of the policy (HM Treasury, 2007^[66]). The costs and benefits of each policy option must be assessed relative to the counterfactual. Each cost and benefit is expressed in monetary terms where possible, to enable their comparison.

As mentioned in the previous section, CBA can be applied at any point of the policy cycle, for both ex-ante, in-itinere and ex-post appraisals. It is particularly useful for comparing policies with the same goal or interventions that address different factors.¹ However, the quality of results depends heavily on available data and the ability to monetise impacts. CBA may also face challenges in monetising variables that are hard or difficult to quantify, such as environmental externalities (Jamison, 2006^[74]). When it is not practical or possible to apply a monetary value, costs and benefits should still be included qualitatively. Because of the qualitative nature of some CBA indicators, and the difficulty to ensure the assessment is comprehensive of all possible costs and benefits, it is particularly important to follow rigorous guidelines on how to perform CBA to ensure consistency and transparency of the assumptions used to perform the analysis.²

TABLE 5.6. Case study: Ex-ante CBA assessment of an enterprise loan scheme in Ireland³

Policy Issue	The Irish Department of Enterprise, Trade and Employment (DETE) uses the Economic Appraisal Mode (EAM) framework to evaluate the effectiveness of state-backed loan schemes for the enterprise sector. The EAM framework assesses the link between the loan scheme and the performance of beneficiaries and the wider economy, the impact on credit availability to firms, and the costs and benefits of the scheme. This framework ensures the best use of economic resources and is consistent with the Irish Public Spending Code. DETE has conducted several ex-ante assessments of proposed loan schemes (e.g., Brexit Loan Scheme; COVID-19 Working Capital Scheme; and Future Growth Loan Scheme), including a new long-term lending scheme for SMEs (including farmers, fishers, and small mid-capital companies) in 2022, which was supported by an ex-ante CBA using the EAM.
Methodology	The EAM involves estimating the anticipated costs and benefits of the proposed project (with adjustments for displacement and deadweight) and calculating the discounted present value of the net benefit of the project over a period of seven years to 10 years. This exercise requires specific data sets, including: administrative data collected through loan applications; data available through the Irish Central Statistics Office from the Business Register; the Annual Services Inquiry; firm data available through commercial databases and survey data directly collected for purpose of evaluation. Collecting the appropriate data can be a challenge, especially for this type of ex-ante evaluation where the required data are not always available. Therefore, this exercise was also informed by international and national examples of similar loan schemes.
Results	The EAM provided a CBA for the proposed new loan scheme aiming at financing SMEs to assist them in expanding and growing. The outcome of the EAM indicated strong and positive economic benefits from the new scheme that supports the business case for the scheme over a timeframe of 10 years. The business case was presented to the Department of Public Expenditure and Reform and the required budget for the new proposed long term loan scheme was approved.

Synthesis methods: systematic reviews and meta-analysis

Systematic reviews and meta-analysis are a qualitative methodology that can be used to estimate the impact of an intervention. They involve systematically analysing and synthesising existing evidence on the effectiveness of different approaches with respect to the desired outcome variable. Systematic reviews are useful for identifying potential impacts of policies based on similar policies implemented in different contexts. They provide a comprehensive overview using explicit processes that are replicable and open to scrutiny by external parties.

The approach can include two phases:

- The “**Systematic Review**” is a qualitative study of studies. It is a summary of the existing literature that uses explicit and reproducible methods to systematically search, critically appraise the evidence on a specific topic, and synthesises the results of different studies.

1. It is, thus, useful also in case the policymaker is interested in limiting duplication and dispersion of policy interventions that may limit their efficiency and effectiveness. This is the goal of the ArCo methodology implemented by the Colombian Government since 2020. ArCo maps policy instruments that seek to boost productivity, competitiveness, and innovation, to identify synergies among them to group initiatives and achieve greater impact, mitigating duplication of efforts (Departamento Nacional de Planeación, 2021^[84]).

2. See for example the guidelines developed by the Danish Ministry of Finance, accessible on the Ministry's website: <https://en.fm.dk/>

3. This case study was presented by DETE experts at the Second Study Visit of the project on 27 October 2022.

- A **meta-analysis** is the statistical analysis of the results of a group of individual studies in order to integrate the findings into a single estimate. Whereas the “systematic review” has the purpose of collecting, reviewing, and presenting all available evidence, the “meta-analysis” refers to the statistical technique involved in extracting and combining data to produce a summary estimate of the impact of a policy.

The method presents some limitations. Biases can occur in selecting and assessing literature, and conducting a full systematic review is a resource-intensive process that can be challenging for non-academic researchers. Systematic reviews with narrowly defined questions only provide specific answers to specific questions, so it is crucial to identify targeted criteria for selecting reviewed articles and to be aware of the sources and scope of the literature reviewed. In addition, meta-analysis can be hampered by methodological diversity in different studies

TABLE 5.7. Case study: A systematic review of sustainable practices for businesses

Policy Issue	Climate change is making the need to adapt business practices to a low-carbon economy. However, sustainability is sometimes associated with higher costs. This case study presents a systematic review of the literature on the relationship between sustainability and competitiveness (Lopez-Torres, 2021 ^[75]).
Methodology	Researchers conducted a systematic review of 182 scientific articles at the intersection of sustainability and competitiveness of firms. Following specific selection criteria, 51 articles were selected for review. The review was organised in two phases: descriptive analysis and content analysis. The selected papers were categorised according to their publication timing, topic area, type of journals, and methodology. The research also assessed how the literature interprets the concepts of sustainability and competitiveness, which sustainable strategies are most often discussed, and which models are used to measure sustainability. It was observed that more papers on the topic adopted a quantitative methodology than a qualitative or conceptual approach.
Results	The case study concluded that the nexus between sustainability and competitiveness is still understudied. In particular, the review identified a need for more mixed research methodologies on sustainability for competitiveness in firms and experimental and design research. It also highlighted a need to establish measurements of the sustainability impact on a firm's competitiveness.

Source: (Lopez-Torres, 2021^[75]).

In-itinere impact evaluation

In-itinere impact evaluations are conducted while the evaluated policy is still being implemented. These assessments do not rely on a specific quantification method, but rather may use a range of both experimental and quasi-experimental designs which can also be used for ex-ante and ex-post evaluations. The experimental designs (RCT) and quasi-experimental designs are examined under ex-ante and ex-post evaluation. This part looks at pilot evaluations that could be a specific approach of in-itinere evaluation.

Pilot interventions only apply to a fraction of the potential beneficiaries. The underlying logic of the pilot is that something is being “tested”, which makes it easier to introduce evaluative considerations into the intervention design. In these cases, non- and quasi-experimental designs can be an alternative. For example, a pilot intervention that is implemented only in certain geographical areas, because the intervention focuses on these regions, can be evaluated with quasi-experimental designs (e.g., matching or difference in difference models that will be explained later) that use the non-pilot areas to build comparison groups.

In all cases, it is necessary to plan good data collection before and after the intervention and to collect data for both potential beneficiaries and non-beneficiaries, to increase the possibilities of obtaining credible impact estimates through non-experimental methods. Indeed, in-itinere evaluations have been criticised for their lack of external validity, as results from small-scale pilot interventions may not be applicable to other contexts. Additionally, such evaluations may not have sufficient time to observe the effects of an intervention, particularly for policies that require years to generate effects. The use of a small sample size also poses a challenge to accurately assess statistically significant differences between treatment and control groups. To mitigate these limitations, evaluators can conduct multi-site evaluations to analyse the variability of impact results across contexts. Supplementing in-itinere evaluations with robust ex-post analyses can help to verify or adjust preliminary findings.

TABLE 5.8. Case study: Evaluating in-itinere Firms' growth patterns in the Netherlands

Policy Issue	<p>Declining business dynamism is a worldwide phenomenon and facilitating startup growth is a major policy goal as a countervailing act. Generally, policies and research about startup ecosystems are mainly based on privately owned data sets (e.g., Crunchbase, Dealroom). However, it is not always clear if commercial databases give complete picture, with the consequence of results and implications being possibly flawed.</p> <p>The aim of this research (El-Dardiry and Vogt, 2022^[76]) was to create a better understanding of start-ups and growth of firms in the Netherlands, exploring questions such as: How useful is an investor database for scientific research? How many High Growth Firms (HGF) are captured in an investor database? Do the growth dynamics differ between 'regular' firms and those in the database?</p> <p>To answer this, the evaluators investigated the general validity of an investor start-up database in studying high growth firms and growth persistence. The analysis was conducted by the Bureau for Economic Policy Analysis (CPB), with a research team composed of two employees and one intern working on the project. Other institutions involved in the study included the Dutch Ministry of Economic Affairs, the Ministry of Finance, and TechLeap, a non-profit publicly funded organisation focusing on start-ups.</p>
Methodology	<p>The researchers started by collecting and building a representative data set to be used later for counterfactual analyses. The data, which was partly hand collected, was merged with full population micro-data on all firms in the period 2010-18. Data collection started at the end of 2019 and was delayed because of the pandemic, with first results presented at the beginning of 2021 and project end in 2021.</p> <p>Addressing the importance of selection effects in the available investor start-up database was a key methodological challenge.</p>
Results	<p>At a practical level, the study contributed to expand the toolkit to study start-ups in the Netherlands, by creating a database which can be used for future research, matching Dutch administrative firm level panel data with this start-up database.</p> <p>The main findings of the study shed light on growth patterns of Dutch firms and start-ups and the selectivity of a commercial database. A strong selection bias was detected in the database. First, although most high growth firms in the Dutch economy do not appear in the studied commercial database, included firms do have a higher probability of being a high growth firm. Second, growth dynamics differ between database firms and other entrants. In contrast to regular Dutch firms and previous findings in the literature, start-ups in the database show a strikingly persistent growth pattern: their growth phases are prolonged.</p> <p>The study concluded that the use of commercial start-up databases for policy-making can complement more traditional data sources, but requires very careful interpretation due to unclear selection in the database. Policies focusing only on these databases might miss a big chunk of firms.</p>

Source: (El-Dardiry and Vogt, 2022^[76]).

Ex-post evaluation

It is often difficult to randomly assign treatment and control groups for ex-post evaluations. Accordingly, statistical and econometric techniques have to be used to estimate the policy impact while minimising selection bias. These methods, commonly defined as quasi-experimental, are examined in the first part of this section, looking in particular at matching, difference-in-difference, instrumental variable and regression discontinuity methods. The case studies presented below focus on the analysis of the impact of R&D support on firm performance. A key challenge for these evaluations is to assess the additionality effect of the support provided to boost R&D investment. In this respect, it is particularly important to collect data on the amount of public support for a meaningful evaluation. If these data are not accessible, it should be explicitly acknowledged that the data only allow to test the combined effect of public and private expenditures. Finally, the rapid impact evaluation method is presented at the end. While this method is not based on an empirical causal approach, it serves as a useful example of the use of ex-post qualitative evaluation.

Matching method

In cases where there are observed characteristics influencing whether individuals are beneficiaries, the estimation of the policy impact could be biased. The solution is to statistically control for the observed characteristics of the individuals that influence participation in the intervention. This approximation allows the evaluator to address the endogeneity problem, and estimate the effect of the intervention. Matching methods are used to estimate the impact of interventions in cases where RCTs are not appropriate. The estimated impact is the average effect of the treatment on all those treated, rather than a marginal impact on a small subset of the treated group. However, matching can only be used if there is a clear and detailed understanding of the eligibility criteria, and the recipient and non-recipient groups have a number of members with similar scores.

Propensity Score Matching (PSM) is a statistical technique that enables evaluators to construct a counterfactual group to estimate the impact of an intervention. The treatment observations are matched with observations in the control group based on their similarity in the probability of being treated (or their propensity score). This is calculated using observable characteristics that determine the likelihood of participation and varies between 0 and 1 (where 1 is 100% likelihood to be treated). By comparing the outcomes of interest between the two matched groups, an impact estimate can be calculated.

Matching requires outcome data for participants and non-participants, as well as data on effective participation in the intervention and baseline characteristics. Propensity Score Matching is only based on observable characteristics, and impact estimates will be biased if there are unobservable factors that affect treatment and outcomes. Matching can only be done on pre-intervention characteristics, which need to be time-invariant or collected beforehand. As such, the matching variables should be selected before the time of assignment as otherwise matching variables could be affected by the treatment itself.

TABLE 5.9. Case study: evaluating the effect of direct public support for innovation in the UK

Policy Issue	In 2014, The UK Department for Business, Innovation and Skill (BIS) commissioned an evaluation (BIS, 2014 _[77]), to study the impact of its innovation policies. The study focused on direct financial support and cooperation with the public sector. Significant financial support was defined as over 20% of in-house R&D spending coming directly from the government.
Methodology	This analysis estimated the impact of grants for R&D on innovation by assessing the impact of direct innovation support, examining the interaction between financial support and cooperation with universities and PSRE, and constructing a counterfactual by using Propensity Score Matching. The dataset was constructed by linking two datasets for the Office of National Statistics (ONS): BERD data and the UK Innovation Survey (UKIS). The evaluators used the Propensity Score Matching approach to construct a quasi-experimental dataset, which allowed them to match grant-holders with highly similar non-grant-holders and use the outcome for these matched 'untreated' firms as a proxy for the outcome that would have been observed for the 'treated' firms had they not participated in the program.
Results	<p>The evaluation provided evidence of a positive impact that direct UK public policy interventions play in terms of generating additional innovation for both SMEs and large firms.</p> <p>In terms of the impact of public funding on private funding for innovation, the analysis found no evidence of deadweight/crowding out. Indeed, the opposite is observed, with public funding crowding in around 30 per cent more private funding (in addition to the public funding provided) over the short term.</p> <p>Lastly, the analysis could not draw any serious conclusions about the impacts on turnover, employment and productivity over the three-year period examined, as a longer time series would have been needed.</p>

Source: (BIS, 2014_[77]).

Difference in difference

The diff-in-diff method is useful for comparing before/after scenarios in situations like the introduction of a new policy or policy changes. In the difference-in-difference method, also known as diff-in-diff, the effect of a treatment or policy is estimated by comparing the pre- and post-treatment differences in the outcome in the treatment and control groups. This method can be appropriate when the selection bias depends on unobserved variables, as it allows unbiased estimates of the impact of the treatment.

However, it has limitations and risks. The quality of the data used is strongly linked to the quality of the method. The method requires baseline and follow-up data on outcomes and other characteristics for both participants and non-participants. The outcomes need to be replicable over time, and the measurement should be independent of the policy's existence. If the effect is expected to manifest after a few years, the results may be contaminated by other events. The method assumes that the two groups would have had a similar trend in the absence of the treatment, and this assumption can be tested by examining common trends before the intervention.

TABLE 5.10. Case study: evaluating the R&D Tax Credit in Ireland

Policy Issue	<p>In 2004, Ireland introduced an R&D tax credit to increase business R&D in the country. The Government’s Innovation Strategy aimed to achieve the EU 2020 target of increasing R&D intensity in Ireland to 2.5 per cent of GNP by 2020. Initially, the scheme employed an incremental R&D expenditure since the first year eligible for the credit (2003). Over 2012 to 2015, the scheme evolved into a full-volume scheme, meaning that all R&D expenditure became eligible for the credit. The credit was made available to all firms liable to Irish tax that undertake R&D activities in the European Economic Area.</p> <p>The R&D tax credit formed part of Ireland’s corporation tax “offering”, aimed at attracting jobs and investment into Ireland and developing a strong, innovation-driven enterprise sector. The evaluation in the case study investigates whether this tax credit resulted in additional R&D expenditure by firms, meaning R&D that would not have taken place in the absence of the tax credit. This additionality is calculated using a treatment and control group. The evaluation was conducted by the Department of Finance in Ireland and published in 2016 (Irish Department of Finance, 2016_[78]).</p>
Methodology	<p>The evaluation used a panel dataset from the Irish Revenue Commissioners and data from the two main enterprise development agencies to assess the impact of the R&D tax credit in Ireland. The evaluation employed a difference-in-difference estimation through fixed effects regression analysis, exploiting a policy change in 2009 when the credit became a repayable tax credit. The change created two groups for comparison, a treated group and a control group. The identification strategy compared the two groups’ R&D growth in a regression framework, assuming that the difference was due to the financial incentives of the R&D tax credit.</p>
Results	<p>The analysis indicated that the tax credit achieved reasonable additionality. The evaluators estimated that 60% of the R&D conducted by firms since 2009 would not have occurred in the absence of the tax credit. Because the tax credit scheme is a general measure, for which all firms are eligible to apply, the deadweight estimate indicates partial crowding out, i.e., firms replacing their own financing with public financing. Additionally, analysis of the firm characteristics of the R&D tax credit showed that it is mainly older, larger and non-Irish firms who derive financial benefit from the scheme, although it is typically Irish firms who benefit more from the repayable credit element of the scheme.</p> <p>The cost of the tax credit programme reached EUR 553 million in 2014, with outstanding (unused) credits of €592 million in addition.</p>

Source: (Irish Department of Finance, 2016_[78]).

Instrumental Variables

The Instrumental Variable regression estimates the impact of a policy intervention by using a different variable (the instrument) to predict treatment in an econometric analysis. This variable is a factor which influences participation in the treatment, but which is uncorrelated with any characteristics affecting the outcome. If an instrument is found which meets these conditions, an estimate of the impact of the treatment can be derived, which removes the bias created by uncaptured characteristics correlated with the treatment.

The IV methodology is useful when interventions may have been implemented in a biased way, individuals may self-select, or there is time-varying selection bias. However, the main limitations are that finding a valid instrument can be difficult, and the instrument must be strongly correlated with the treatment. The data requirements are similar to other methods like matching and difference-in-difference methods. The derived impact estimate is a Local Average Treatment Effect, which captures the impact of the policy on those who have benefitted from it because of the instrument. Misinterpreting these estimates about the average effect of a policy may result in misleading recommendations. A growing econometric literature discusses how to identify the external validity of Local Average Treatment Effects for the whole population of beneficiaries (Flores and Chen, 2018_[79]).

Regression discontinuity

There are cases where participants are assigned to treatment and control groups based on a cut point of an assignment variable. If a participant is below (above) a certain threshold value, that individual is eligible to receive the intervention, and above (below) that value it is not. In these cases, the regression discontinuity method can be applied to measure the differences (discontinuity) between the treatment and control trends.

Regression discontinuity can be a useful method where randomisation is not possible in the intervention design. Policies assigned through a threshold score are particularly amenable to a regression discontinuity design. However,

TABLE 5.11. Case study: evaluating the impact of R&D subsidies based on the ERDF population-density rule in Finland

Policy Issue	<p>The Finnish government has implemented an R&D support policy aimed at increasing the amount of new R&D projects and support ongoing ones. The program is available to all Finnish firms and is implemented by the Finnish Funding Agency for Technology and Innovation (Tekes). Recipients are selected based on factors such as commercial potential, technological challenge, and the availability of resources. However, the program must comply with EU regulations, which require that the R&D support induces firms to pursue research they would not have pursued without the program.</p> <p>From 2000 to 2005, Tekes granted direct R&D subsidies worth 968 million euros, mostly from the Finnish government or the European Regional Development Fund. The funding was distributed based on districts assigned in accordance with EU criteria, such as population density. However, there was a regional variation in government funding due to Tekes' ability to withdraw larger amounts from the ERDF in areas eligible for higher levels of aid, resulting in substantial variation in the probability of being accepted into the program across areas.</p> <p>The evaluation (Einiö, 2014^[80]) aimed to improve understanding of the potential of R&D subsidies to enhance private innovative efforts in technologically advanced economies.</p>
Methodology	<p>Data used in this evaluation was constructed by linking different administrative datasets to an annually conducted R&D survey, all maintained by Statistics Finland. The ex-post evaluation was conducted by the Finnish government Agency VATT Institute for Economic Research, which operates under the domain of the Ministry of Finance.</p> <p>The primary aim of the evaluation was to assess whether, as a result of becoming a program participant, the firm increased its R&D expenditure. The "treatment group" included the firms that did not receive any R&D support in the year $t - 1$, entered the R&D support program in the year t, and received R&D support in the year $t + 1$. The "control group" included the firms that did not receive any payments from the agency in the consecutive three years $t - 1$, t, and $t + 1$.</p>
Results	<p>The study found that the research grant had a positive impact on winning firms' labor productivity. The probability of receiving an R&D grant was higher in eligible areas for regional aid. This difference provided variation in R&D support program participation, which was used to identify the causal effects of the R&D support on company performance. The IV analysis suggested positive impacts on R&D investment, employment, and sales. However, the IV approach only identified the effect of the R&D support program on firms that changed their participation status due to higher government funding in their region, and the results cannot be interpreted as evidence of the overall effectiveness of the intervention. The effectiveness of the R&D support may vary among participating projects.</p>

Source: (Einiö, 2014^[80])

regression discontinuity also presents limitations, including the need for good data on eligibility cutoff and baseline outcomes. Adequate observations surrounding the cutoff are necessary for statistical power. This method is only applicable to interventions where being a beneficiary strongly depends on a continuous characteristic such as age or number of workers. It estimates the impact of the intervention only for individuals close to the cut-off (local average treatment effect), while the effect may differ for those further away. Econometric methods have been developed to estimate the effects for those further away from the cut-off (Angrist and Rokkanen, 2015^[81]).

Rapid Impact Evaluation

A Rapid Impact Evaluation (RIE) provides a structured way to gather an assessment of an intervention's impact by engaging a number of stakeholders and experts to provide a balanced perspective on the impacts of a policy. Each stakeholder/expert assesses policy outcomes relative to a counterfactual, which is an alternative intervention design or situation, in order to assess the intervention's impact. Different groups of people can be consulted, including stakeholders who affect the intervention or are affected by it (e.g., beneficiaries, policy makers, implementation staff); or external subject-matter experts (e.g. researchers, academics, industry leaders).

RIE is a quick, low-cost evaluation method that relies on expert assessments of an intervention's impact relative to an alternative. RIE is suitable when other assessment methods are challenging or when there is a need for external perspectives. However, RIE has limitations, including potential biases in expert assessments and difficulties in selecting appropriate subject-matter experts and technical advisors. In addition, some contexts may be impractical for an RIE, particularly if a policy is too complex for a compelling counterfactual to be developed or for experts to reasonably assess policy impacts.

TABLE 5.12. Case study: evaluating R&D Subsidies in Italy

Policy Issue	<p>In 2003, the unique “Regional Program for Industrial Research, Innovation and Technological Transfer,” program was implemented in a region of northern Italy (Emilia-Romagna). The program aimed to support industrial research and precompetitive development by firms in the region. Firms were invited to submit proposals for new projects and only those which scored above a certain threshold received the subsidy. The grants covered up to 50 percent of the costs of industrial research projects and 25 percent for precompetitive development projects. The maximum grant per project was EUR 250,000 and the usual duration of the investment from 12 to 24 months. Grants were assigned after a committee of independent experts appointed by the regional government had assessed the projects.</p> <p>The program aimed to assess the impact of such policies to gain a clearer picture of the use of public resources. The evaluation was conducted by the Bank of Italy and published in 2015 (Pavone and Sella, 2015^[82]).</p>
Methodology	<p>The impact of the policy was assessed using balance sheet data to determine if the program enabled outlays that would not have been made without the grant. The program’s effectiveness was assessed using a sharp regression discontinuity design, taking advantage of the program’s local dimension and the large number of participating firms. Overall, the program granted about €93 million to 1246 applicant enterprises in the region of Emilia-Romagna, with each subsidised firm receiving an average of EUR 182,000.</p>
Results	<p>According to the evaluation, the program did not create additional investment. The results do not reject the hypothesis that firms substituted public for privately financed R&D. This overall effect, however, masked substantial heterogeneity in the impact of the program. When the effect of the program was estimated by firm size, the evaluators found that, unlike large firms, small enterprises increased their investment substantially, by on average the same amount of the grant received. The findings were robust to multiple sensitivity checks. As a result, it was implied that more work should be done to better understand all the forces driving the heterogeneous effect of subsidies.</p>

Source: (Pavone and Sella, 2015^[82]).

TABLE 5.13. Case study: evaluating EcoENERGY for Alternative Fuels Program (ecoEAF) in Canada

Policy Issue	<p>Canada’s transportation sector heavily relies on conventional fossil fuels, which contributes to greenhouse gas emissions. To address this, the Natural Gas Use in Transportation Roundtable created the EcoENERGY for Alternative Fuels Program (ecoEAF) program in 2011/12. This program aims to increase the use of natural gas in medium and heavy-duty vehicles and has two pillars: education and outreach and codes and standards.</p> <p>An evaluation was conducted using the RIE approach from 2011-12 to 2014-15 to assess the program’s relevance, effectiveness, efficiency, and economy.</p>
Methodology	<p>The evaluation included a detailed summary of the program, surveys and interviews with program stakeholders, consultations with subject matter experts, a literature review, and analysis of program data. The overall impact of the program was estimated by triangulating and synthesising the data from all lines of evidence.</p>
Results	<p>The study indicated that the ecoEAF program had an early positive impact on the adoption of energy efficient technologies and alternative energy practices and met or exceeded most of its performance targets. Its key activities led to increasing awareness of alternative fuel options and more harmonised codes and standards for alternative fuels, particularly natural gas. The program was highly relevant, but the need has shifted towards requiring support to de-risk investments and early adoption while increasing capacity to sustain markets. The program has ended, but a new program called the ‘Smart Renewables and Electrification Pathways Program (SREPs)’ has been launched to support smart renewable energy and grid modernisation projects that will lower emissions by investing in clean energy technologies such as storage and modernising electricity system operations.</p>

Source: Government of Canada, <https://www.nrcan.gc.ca/maps-tools-publications/publications/minerals-mining-publications/evaluation-report-ecoenergy-alternative-fuels-program/18646#s1>

References

- Albanese, G. et al. (2021), “Nuove evidenze sugli aiuti alle imprese in Italia”, *Banca d’Italia Occasional Paper* 640, <https://doi.org/10.32057/0.QEF.2021/0640>. [83]
- Angrist, J. and M. Rokkanen (2015), “Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away from the Cutoff”, *Journal of the American Statistical Association*, Vol. 110/512, pp. 1331-1344, <https://doi.org/doi.org/10.1080/01621459.2015.1012259>. [81]
- Bender, S. et al. (2019), “The Annodata Framework: Putting FAIR data into practice”, *Technical Report* 2019-03, Deutsche Bundesbank, Research Data and Service Centre. [53]
- Bender, S., B. Hausstein and C. Hirsch (2018), “An introduction to INEXDA’s metadata schema”, IFC Bulletins chapters, in: *Bank for International Settlements (ed.), Are post-crisis statistical initiatives completed?*, volume 49, Bank for International Settlements. [38]
- BIS (2014), “Estimating The Effect Of UK Direct Public Support for Innovation”, *Bis Analysis Paper number* 04, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/369650/bis-14-1168-estimating-the-effect-of-uk-direct-public-support-for-innovation-bis-analysis-paper-number-04.pdf. [77]
- BMWK (n.d.), *Evaluations*, <https://www.bmwk.de/Navigation/DE/Service/Evaluationen/evaluationen.html> (accessed on 10 June 2022). [29]
- Calvino, F. et al. (2022), “Closing the Italian digital gap: The role of skills, intangibles and policies”, *OECD Science, Technology and Industry Policy Papers*, OECD Publishing, Paris 126, <https://doi.org/10.1787/e33c281e-en>. [37]
- Ciani, E., M. Gallo and Z. Rotondi (2020), “Public credit guarantees and financial additionalities across SME risk classes”, *Banca d’Italia Temi di discussione* Nr. 1265, <https://www.bancaditalia.it/pubblicazioni/temi-discussione/2020/2020-1265/index.html?com.dotmarketing.htmlpage.language=1>. [16]
- Cole, S. et al. (2020), “Handbook on Using Administrative Data for research and evidence-based policy”, <https://doi.org/10.31485/admindatahandbook.1.0>. [56]
- Crato, N. and P. Paruolo (2019), *Data-Driven Policy Impact Evaluation: How Access to Microdata is Transforming Policy Design*, Springer Open, <https://doi.org/10.1007/978-3-319-78461-8>. [36]
- Cribillers, C. (2022), “A solid milestone in the reattempt for Spanish evaluation, European Evaluation Society”, <https://europeanevaluation.org/2022/05/19/a-solid-milestone-in-the-reattempt-for-spanish-evaluation/>. [9]
- DAGL (2020), *Relazione al Parlamento sullo stato di applicazione dell’analisi d’impatto della regolamentazione*, Presidenza del Consiglio dei Ministri, https://presidenza.governo.it/DAGL/uff_studi/RelazioneAnnuale2020.pdf (accessed on 28 February 2022). [6]
- Danish Business Authority (2013), “Nudging business policy, making it easy to do the right thing”, <http://www.copenhageneconomics.com/dyn/resources/Publication/publicationPDF/3/243/0/Nudging%20Business%20Policy.pdf>. [72]
- de Blasio, G. et al. (2018), “Public Guarantees to SME Borrowing. A RDD Evaluation”, *Journal of Banking and Finance*, Vol. 96, pp. 73-86, <https://doi.org/10.1016/j.jbankfin.2018.08.003>. [17]
- Departamento Nacional de Planeación (2021), *Metodología de Articulación para la Competitividad*, ArCo, <https://www.dnp.gov.co>. [84]
- Department of Business, Energy & Industrial Strategy (2020), “BEIS Monitoring and Evaluation Framework” 2020/016, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/947722/beis-monitoring-evaluation-framework.pdf (accessed on 11 April 2022). [31]

References

- Desai, T. (2014), “Recommendations for Improving Research Access to Potentially Disclosive Data”, *Welsh Government*. [57]
- Domeyer, A. et al. (2021), *Government data management for the digital age*, <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/government-data-management-for-the-digital-age> (accessed on 09.2022). [42]
- EC (2014), *Monitoring and Evaluation of European Cohesion Policy: Guidance document on ex-ante evaluation*, https://ec.europa.eu/regional_policy/sources/docoffic/2014/working/ex_ante_en.pdf. [15]
- EC (2013), “Evalsed Sourcebook-Method and Techniques”, *European Commission*, http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm. [2]
- Einiö, E. (2014), “R&D SUBSIDIES AND COMPANY PERFORMANCE: EVIDENCE FROM GEOGRAPHIC VARIATION IN GOVERNMENT FUNDING BASED ON THE ERDF POPULATION-DENSITY RULE.”, *The Review of Economics and Statistics*, 96(4), 710–728. <http://www.jstor.org/stable/43554950>. [80]
- El-Dardiry, R. and B. Vogt (2022), *How Far Do Gazelles Run? Growth Patterns of Regular Firms, High Growth Firms and Startups.*, No. 432. CPB Netherlands Bureau for Economic Policy Analysis. [76]
- Finaldi Russo, P., S. Magri and C. Rampazzi (2016), “Innovative start-ups in Italy: they special features and the effects of the 2012 law”, *Questioni di Economia e Finanza*, Vol. 339, https://www.bancaditalia.it/pubblicazioni/qef/2016-0339/QEF_339_16.pdf. [20]
- Flores, C. and X. Chen (2018), “Average Treatment Effect Bounds with an Instrumental Variable: Theory and Practice”, *Springer Singapore*. <https://doi.org/10.1007/978-981-13-2017-0>. [79]
- Gobierno de España (2022), *El Gobierno aprueba el anteproyecto de Ley de Evaluación de Políticas Públicas para fortalecer el proceso de análisis y la eficacia de las medidas adoptadas*, Retrieved from: <https://planderecuperacion.gob.es/noticias/el-gobierno-aprueba-el-anteproyecto-de-ley-de-evaluacion-de-politicas-publicas-para-fortalecer>. [7]
- Hawley, J. (2020), *Ohio and the Longitudinal Data Archive: Mutually Beneficial Partnerships Between State Government and Researchers*, Abdul Latif Jameel Poverty Action Lab. [63]
- HM Treasury (2020), *Magenta Book: Central Government guidance on evaluation*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf. [67]
- HM Treasury (2018), *The Green Book: Central Government Guidance on Appraisal and Evaluation*, London Crown, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063330/Green_Book_2022.pdf. [68]
- HM Treasury (2007), “The Magenta Book: guidance notes for policy evaluation and analysis.”, *Magenta Book Background Papers*, https://dera.ioe.ac.uk/10521/1/complete_Magenta_tcm6-8611.pdf. [66]
- IBM (2020), *Data Warehouse*, <https://www.ibm.com/cloud/learn/data-warehouse#toc-what-is-a--1NvW7Mkj>. [46]
- IBM (2019), *IaaS (Infrastructure-as-a-Service)*, <https://www.ibm.com/cloud/learn/iaas#toc-what-is-ia-k7Negcdh>. [50]
- INEXDA (2020), *Working group on data access*, Deutsche Bundesbank, Research. [52]
- INPS (2021), *Programma Visitinps scholars*, <https://www.inps.it/dati-ricerche-e-bilanci/attivita-di-ricerca/programma-visitinps-scholars> (accessed on 29 June 2022). [60]
- Invitalia (2022), *Programmi operativi: PON Imprese e Competitività*, <https://www.invitalia.it/cosa-facciamo/affianchiamo-la-pa-per-gestire-i-fondi-europei-e-nazionali/programmi-operativi/pon-impres-e-competitivita>. [14]
- Invitalia, MET and CSIL (2020), *Affidamento dei servizi di valutazione del regime di aiuti attuato mediante l'intervento agevolativo “Beni strumentali – Nuova Sabatini”*: Rapporto di valutazione, https://www.mise.gov.it/images/stories/documenti/Report_finale_2020_NS.pdf. [21]

- Irish Department of Finance (2016), “Economic Evaluation of the R&D Tax Credit”, *An Roinn Airgeadais*, <https://igees.gov.ie/wp-content/uploads/2014/01/R-and-D-Credit-Evaluation-2016.pdf>. [78]
- Jamison, D. (2006), “Cost-Effectiveness Analysis”, 2006. *Priorities in Health*, ed., 39-58. New York: Oxford University Press. [74]
- Jones, C. et al. (2022), “Building on Aotearoa New Zealand’s Integrated Data Infrastructure”, *Harvard Data Science Review* 4.2, <https://doi.org/10.1162/99608f92.d203ae45>. [41]
- Jones, S. et al. (2019), “Risks and rewards of cloud computing in the UK public sector: A reflection on three Organisational case studies”, *Information Systems Frontiers* 21, pp. 359–382, <https://doi.org/10.1007/s10796-017-9756-0>. [48]
- Ladley, J. (2019), *Data Governance. How to Design, Deploy, and Sustain an Effective Data Governance Program*, Academic Press. [33]
- Langedijk, S., I. Vollbracht and P. Paruolo (2019), *The Potential of Administrative Microdata for Better Policy-Making in Europe*, in “The Power of Microdata: An Introduction: How Access to Microdata is Transforming Policy Design”. [39]
- Lopez-Torres, G. (2021), “Sustainability for competitiveness in firms – a systematic literature review”, *Measuring Business Excellence*, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/MBE-02-2021-0023>. [75]
- Maillet, D. and J. Ted McDonald (2020), *New Brunswick Institute for Research, Data and Training: A Ten-Year Partnership Between Government and Academia*. [61]
- Manaresi, F. (2021), “Imprese”, in de Blasio, G.; Nicita, A.; Pammolli, F. (eds), *Evidence-based Policy! Ovvero perché politiche pubbliche basate sull’evidenza empirica rendono migliore l’Italia*, Il Mulino. [18]
- Menon, C. et al. (2018), “The evaluation of the Italian “Start-up Act””, *OECD Science, Technology and Industry Policy Papers*, Vol. No. 54, <https://doi.org/10.1787/02ab0eb7-en>. [3]
- Microsoft (2022), *Cloud computing terms*, <https://azure.microsoft.com/en-gb/resources/cloud-computing-dictionary/>. [51]
- MIMIT (2022), *Relazione annuale sullo stato di attuazione della normativa a sostegno delle startup e delle PMI innovative*, <https://www.mise.gov.it/index.php/it/impresa/competitivita-e-nuove-imprese/start-up-innovative/relazione-annuale-e-rapporti-periodici>. [12]
- MIMIT (2021), *Relazione annuale sullo stato di attuazione della normative a sostegno delle startup e delle PMI innovative*, https://www.mise.gov.it/images/stories/documenti/RELAZIONE_ANNUALE_2021.pdf. [11]
- MIMIT (2021), *Relazione sugli interventi di sostegno alle attività economiche e produttive anno 2021*, https://www.mise.gov.it/images/stories/documenti/RELAZIONE266_WEB2021.pdf. [10]
- MIMIT (n.d.), *PNRR – Transizione 4.0*, <https://www.mise.gov.it/index.php/it/pnrr/progetti-pnrr/pnrr-transizione-4-0>. [22]
- MIMIT and ISTAT (2018), *Startup Survey 2016: The First Survey of Innovative Startups in Italy*, ISTAT, https://www.istat.it/it/files//2018/06/ebook_STARTUP_SURVEY_english_version.pdf. [19]
- Ministerio de Hacienda y Función Pública (2022), “El Gobierno aprueba el anteproyecto de Ley de Evaluación de Políticas Públicas para fortalecer el proceso de análisis y la eficacia de las medidas adoptadas (Nota de prensa)”, Accessible at: <https://www.hacienda.gob.es/Documentacion/Publico/GabineteMinistro/Notas%20Prensa/2022/CONSEJO-DE-MINISTROS/19-04-22-NP-CM-ANTEPROYECTO-LEY-EVALUACION-POLITICAS-PUBLICAS.pdf>. [8]
- Müller, D. and J. Möller (2019), “Giving the International Scientific Community Access to German Labor Market Data: A Success Story”, In: Crato, N., Paruolo, P. (eds) *Data-Driven Policy Impact Evaluation*. Springer, Cham. https://doi.org/10.1007/978-3-319-78461-8_7. [58]

References

- Müller, D. and P. vom Berge (2020), *Institute for Employment Research, Germany: International Access to Labour Market Data*, in “Using administrative data for research and evidence-based policy”. [59]
- Müller, M., M. Kudic and A. Pyka (2017), “Ex-Ante Evaluation of Investments in Knowledge, Learning, and Innovation: An Application Example”, TATuP, Available from: <https://www.tatup.de/index.php/tatup/article/view/67>. [73]
- NIIS (2022), *Nordic Institute for Interoperability Solutions (NIIS)*, <https://www.niis.org/>. [45]
- Nordic Council of Ministers (2019), “Policy instrument evaluation: A tool for increasing efficiency in environmental policy”, *Nordic Co-operation*. Accessible at: <https://www.norden.org/en/publication/policy-instrument-evaluation>. [70]
- OECD (2021), “Applying Evaluation Criteria Thoughtfully”, *OECD Publishing*, Paris, <https://doi.org/10.1787/543e84ed-en>. [64]
- OECD (2020), *Building Capacity for Evidence-Informed Policy-Making: Lessons from Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://dx.doi.org/10.1787/86331250-en>. [30]
- OECD (2020), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://dx.doi.org/10.1787/89b1577d-en>. [23]
- OECD (2020), *Mobilising Evidence for Good Governance: Taking Stock of Principles and Standards for Policy Design, Implementation and Evaluation*, OECD Publishing, Paris, <https://doi.org/10.1787/3f6f736b-en>. [65]
- OECD (2019), “The Path to Becoming a Data-Driven Public Sector”, *OECD Digital Government Studies*, OECD Publishing, Paris, <https://doi.org/10.1787/059814a7-en>. [43]
- OECD (2019), “Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit”, OECD Publishing, Paris, <https://doi.org/10.1787/9ea76a8f-en>. [71]
- OECD (2018), “Trade in Counterfeit Goods and the Italian Economy: Protecting Italy’s intellectual property”, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264302426-en>. [4]
- OECD (2016), *Open Government: The Global Context and the Way Forward*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264268104-en>. [25]
- OECD (2014), “Cloud Computing: The Concept, Impacts and the Role of Government Policy”, *OECD Digital Economy Papers*, OECD Publishing, Paris No. 240, <https://doi.org/10.1787/5jxzf4lcc7f5-en>. [49]
- OECD (2014), “Expert Group for international collaboration on microdata access”, <https://www.oecd.org/sdd/microdata.htm>. [54]
- OECD (n.d.), *Regulatory impact assessment*, <https://www.oecd.org/regreform/regulatory-policy/ria.htm> (accessed on 13 September 2022). [5]
- Oliver, K. et al. (2015), “Identifying public health policymakers’ sources of information: comparing survey and network analyses”, *The European Journal of Public Health*, p. ckv083, <https://doi.org/10.1093/eurpub/ckv083>. [26]
- Patton, M. (1978), “Utilization-focused evaluations”, Beverly Hills: Sage, 1, 978. [24]
- Pavone, S. and E. Sella (2015), “Politiche di sostegno al settore agroindustriale in Piemonte: una valutazione controfattuale [Supporting agro-food enterprises in Piedmont: a counterfactual evaluation]”, *IRCrES Working Paper 201501*. [82]
- PCM (n.a.), *Italia Domani: Piano Nazionale di Ripresa e Resilience: Catalogo Open Data*, <https://italiadomani.gov.it/it/catalogo-open-data.html?orderby=%40jcr%3Acontent%2FobservationDateInEvidence&sort=desc>. [13]
- Presidenza del Consiglio dei Ministri (2018), “Guida all’analisi e alla verifica dell’impatto della regolamentazione”, Accessible at: <https://presidenza.governo.it/DAGL/pdf/GuidaAIR.pdf>. [1]
- Shlomo, N. (2020), *Overview of Data Linkage Methods for Policy Design and Evaluations*. [40]

- Silberman, R. and M. Alkhoury (2021), *Transnational Access to Confidential Microdata*. [55]
- Tillväxtanalys (2022), *Analysis and evaluation plan for the period 2022-2024*, https://www.tillvaxtanalys.se/download/18.4f55b99217f698d4c195fd64/1649402734649/PM_2022_02_01_Tillv%C3%A4xtanalys%20AU_plan_2022.pdf (accessed on 9 June 2022). [28]
- Tillväxtanalys (n.d.), *Vetenskapligt råd*, <https://www.tillvaxtanalys.se/om-oss/var-verksamhet/vetenskapligt-rad.html> (accessed on 9 June 2022). [32]
- Ubaldi, B. (2013), “Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives”, *OECD Working Papers on Public Governance*, OECD Publishing, Paris 22, <https://doi.org/10.1787/5k46bj4f03s7-en>. [34]
- Umbach, G., C. Guidi and M. Russo (2018), *Evidence-Based Policy-Making: from data to decision-making*, European University Institute (EUI), <https://doi.org/10.2870/869841>. [62]
- Van Ooijen, C. (2019), “A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance”, *OECD Working Papers on Public Governance*, OECD Publishing, Paris. 33, <https://doi.org/10.1787/09ab162c-en>. [35]
- Westhorp, G. (2014), “Realist Evaluation: AN Introduction.”, Methods Lab Overseas Development Institute London. [69]
- World Bank (2022), *Cloud Services Advance Digital Transformation for Governments*, <https://www.worldbank.org/en/news/feature/2022/06/07/cloud-services-advance-digital-transformation-for-governments>. [47]
- X-Road (2022), *The business registers of Estonia and Finland start cross-border interoperability*, <https://x-road.global/case-study-the-business-registers-of-estonia-and-finland>. [44]

OECD countries are placing increasingly greater emphasis on policies aimed at supporting business and economic development. The design and implementation of these policies need to be informed by good evidence on their expected and actual impacts to understand whether support has been effective and ensure that resources and interventions are deployed effectively and sustainably.

In recent years, the Ministry of Enterprises and Made in Italy has made important efforts to promote a stronger culture of evaluation and monitoring within the institution. The recent establishment of a new Analytical Unit (Centro Studi) responsible for conducting ex-ante, in-itinere and ex-post policy evaluations marks an important step towards enhancing the Ministry's capacity to produce and use impact evaluations.

This report presents the result of work aimed at supporting the development of the Centro Studi and the use of evidence in the Ministry. It draws on best international practices to provide recommendations, guidance and methodologies to strengthen the Centro Studi's governance, data, evaluation capacity and ability to inform the design of the Ministry's economic policies, including on industrial policies, innovation, entrepreneurship and productivity.

The work underlying this report was conducted by an interdisciplinary OECD team bringing together the OECD Economics Department, the Public Governance Directorate and the Directorate for Science, Technology and Innovation.

The project was funded by the European Union via the Technical Support Instrument, and implemented by the OECD, in cooperation with the Directorate-General for Structural Reform Support of the European Commission.

