

5

An improved monitoring and evaluation framework for social inclusion in Spain

This chapter discusses the advantages of developing a monitoring and evaluation (M&E) framework for Spain's new social inclusion model for the beneficiaries of the national minimum income scheme (Ingreso Mínimo Vital, IMV). The chapter outlines a high-level framework and examines the human resources necessary for conducting M&E, both in terms of staff requirements and the use of external resources. It also discusses the data requirements. Ideally, the framework should be developed and implemented in co-operation with relevant stakeholders. The chapter also provides specific examples for implementation regarding both the national-level framework and regional-level pilots.

5.1. A general framework for the monitoring and evaluation of Spain's new social inclusion model

This section aims to provide insights for Spain in its efforts to introduce a monitoring and evaluation (M&E) framework for a new social inclusion model (SIM) for the national minimum income (Ingreso Mínimo Vital, IMV) beneficiaries, focussing on the national-level dimensions of such a framework. It first discusses the concept of M&E frameworks from a theoretical perspective and provides some background on institutional set-ups and practices of implementing similar frameworks in Spain. Next, the chapter outlines previous evidence generation on IMV beneficiaries and their social and labour market integration. It proposes a new systematic approach to monitor and evaluate the SIM of IMV beneficiaries, putting forward considerations for objectives, indicators, methodologies, human resources, co-operation requirements and data needs for such a framework. These various dimensions aim to guide Spain in finetuning a more specific action plan to develop the new M&E framework, taking into account the feasibility of co-operating with the Autonomous Communities of Spain (Comunidades Autónomas de España, hereafter “AACC”) and other ministries in this process and linking the M&E framework into a broader accountability concept and evidence-based policy making cycle.

5.1.1. Objectives and premises

This subsection defines M&E activities and explains their objectives, importance and general design. Next, the section looks at the practices and institutional set-up of M&E frameworks in Spain, focusing above all on the policy fields of social security, social services and employment.

What M&E frameworks are

The policies and programmes developed under the new SIM must be introduced in a setting where transparency and accountability are promoted and effectiveness and efficiency are analysed. Evidence-based policy making needs to be systematic and integrated into a coherent cycle that includes policy and evaluation design based on available knowledge, implementation, M&E, generation of new knowledge/evidence from M&E results, dissemination of that knowledge, and a process of continuous feedback for ongoing improvement based on the evidence generated.

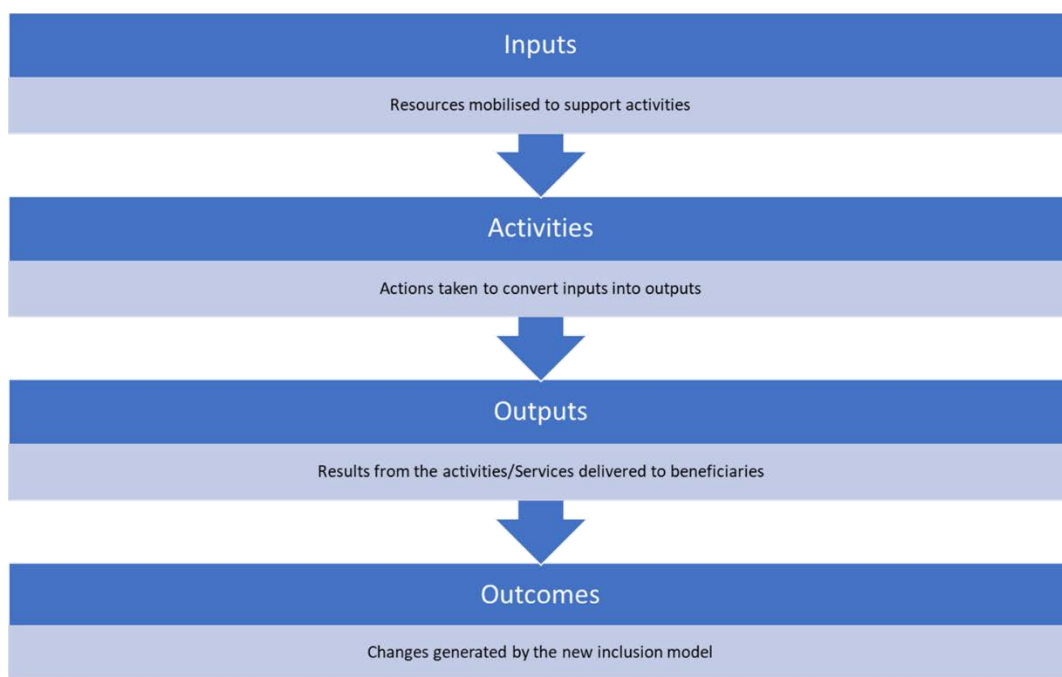
The M&E framework is thus a critical component of the policy-making cycle as it provides objective evidence for policy makers to make informed decisions and adapt policies to achieve better outcomes. The design of the M&E system should be developed simultaneously with the design of the new SIM. It is essential to consider the implications of variables to be included in data collection, data collection mechanisms, staff and skill requirements and possible evaluation methodologies in advance to make the later M&E activities feasible. This way, performance indicators and evaluation criteria are established in advance, enabling immediate and ongoing assessment of the model's execution. The M&E system thus serves as a quasi-real-time feedback mechanism, making it possible to assess the success of the new SIM and to improve its implementation process in a timely manner.

M&E are two complementary pillars for generating evidence to determine whether a programme has met its intended results. Monitoring is an ongoing process of gathering and analysing information about a programme to assess its implementation and development over time while tracking performance against expected results. In contrast, evaluation activities focus on determining how effectively the programme has been implemented, whether it meets its objectives, whether there are causal links between the programme and its results, and whether its benefits outweigh its costs.

The M&E framework uses a theory of change (or intervention logic) to describe how policies or programmes are designed to generate desired outcomes (or impact). The theory of change guides the selection of indicators, data collection methods, and data analysis necessary for identifying and measuring

impact. Often presented as a results chain, it outlines the steps required for a programme to achieve its goals. The results chain typically encompasses four stages: inputs, activities, outputs and outcomes (Gertler et al., 2016^[11]). Inputs refer to the resources required for programme execution. Activities involve specific actions to implement the programme, transforming inputs into outputs. Outputs represent the immediate results of these activities, whereas outcomes are the long-term changes or impacts the programme intends to achieve. Outcomes can be categorised into intermediary outcomes, which support the attainment of ultimate objectives, and final outcomes, which represent the ultimate objectives themselves. The first step in establishing an M&E framework for the new SIM is to identify its results chain (see Figure 5.1). This exercise is also pertinent when planning the pilots of the different elements of the model (see Section 5.2).

Figure 5.1. The results chain framework



Source: OECD internal methodology.

By obtaining credible evidence on the effectiveness of the new SIM through a robust M&E framework, the Spanish Ministry of Inclusion, Social Security, and Migration (MISSM) can effectively convey its significance to the public and policy makers alike, thereby attracting and retaining the necessary resources.

The current framework and practices of evidence generation on social and labour market integration in Spain

The absence of a unified and structured M&E framework that spans national, regional and local levels has resulted in an inadequate system of evidence-based policy making. This leads to a lack of systematic evidence on the efficiency and effectiveness of inclusion policies and an incomplete feedback process. Even when evaluations are conducted, their findings are seldom disseminated and rarely inform policy-making decisions. The implementation of the new Law on the Institutionalisation of Public Policy Evaluation aims to address these existing gaps and has the potential to promote evidence-based policy making in the context of the new SIM.

National-level institutional set-up for M&E

In Spain, Law 27/2022 on the Institutionalisation of Public Policy Evaluation, approved in December 2022, establishes the legal framework for M&E¹ of public policy. The law aims to create a transversal and systematic M&E approach by implementing basic organisational structures and planning mechanisms at the state level. To that aim, the law introduces many key features. It introduces a common set of indicators for all public bodies, departmental co-ordination units to oversee the M&E activities within each ministry, and a requirement for public policies to be evaluated by a team external to the body responsible for implementing the policy. Specific training plans on public policy evaluation will be designed for public employees, and a state agency will be created to oversee evaluation teams and their activities.

The law also establishes two instruments for public policy evaluation: a four-year government Strategic Evaluation Plan at the central state level and a biennial departmental evaluation at the ministry level. Both instruments include *ex ante* and *ex post* evaluation obligations for policies with significant budgetary, economic and social impacts. Institutions responsible for public policies must report on whether they adopt the evaluation report recommendations and justify their decision.

The law is not functional yet, and so far, Spain lacks a common M&E framework. At the national level, the Independent Authority for Fiscal Responsibility (Autoridad Independiente de Responsabilidad Fiscal, AIReF) has extensive and transversal experience in *ex post* evaluation. The Spanish Government periodically commissions AIReF to carry out an in-depth analysis of public spending – the “Spending Reviews” – and other specific evaluations, such as the evaluation of the IMV (AIReF, 2019^[2]). AIReF can also carry out evaluations at the request of the AACC in areas such as healthcare, active labour market policies (ALMPs), education and other public-resource-related areas. Additionally, other ministries and units have performed *ex post* evaluations on an *ad hoc* basis, and the National Agency for Evaluation and Quality (Agencia Estatal de Evaluación y Calidad, AEVAL), which existed until 2017, performed some evaluations of public policy at the national level. In specific public activity sectors, the Sectoral Conferences, which are the bodies in charge of establishing multilateral co-operation between the ministry in charge of that sector at the level of the state and the AACC ministries responsible for the same area, could adopt joint plans to commit to joint actions. In that context, the agreement approving the plans must detail, among others, the mechanisms for M&E.

National-level knowledge generation in the field of social and labour market inclusion before the introduction of the new national minimum income scheme

Regarding inclusion policies, the MISSM does not have a common framework for M&E. The M&E frameworks in Spain are defined only through national strategies focusing on specific populations and objectives. For instance, the National Strategy for Preventing and Fighting Poverty and Social Exclusion 2019-23 (overseen jointly with the Ministry of Social Rights and Agenda 2030, MIDSA) includes an M&E framework that defines the quantitative and qualitative indicators to be monitored and foresees an impact evaluation to retrieve the causal impact of the different actions included in the strategy.

The production of statistics, analysis and evaluations in the fields of labour and social affairs under the Ministry of Labour and Social Economy (MITES), MISSM and MIDSA has largely taken place in silos for different policy areas, such as employment services, social services and social security measures (OECD, 2020^[3]; 2022^[4]). As such, it has not been possible to monitor and evaluate different social and labour market integration pathways comprising different types of support and services. A similar lack of comprehensive cross-policy knowledge tends to be present in most OECD countries (OECD, 2021^[5]).

The focus of data analytics in labour and social affairs in Spain has been mainly on producing descriptive statistics on different policy interventions to disseminate general knowledge on take-up and socio-economic characteristics of beneficiaries of the policies. The tables to disseminate statistics (generally not in interactive format but in PDFs [portable document format] and sometimes Excel) on the

dedicated websites of the relevant ministries and sub-units are supported by analytical reports that aim to summarise and discuss the regular statistics, as well as occasionally present complementary or in-depth statistics, as well as compare with international statistics (e.g. from Eurostat) when feasible. Generally, the reports provide similar information from one period to the next and are simply updated with the newest available data.

Still, the statistics and analyses produced by the many dedicated units under the three main ministries for labour and social affairs in Spain serve some monitoring purposes. The statistics make it possible to observe some progress in the implementation of services and measures over time, particularly related to outputs and, to some degree, inputs.

Nevertheless, the statistics production on most of the measures and services cannot be considered fully-fledged monitoring frameworks for these measures and services, as in most cases, there are no specific lists of monitoring indicators agreed upon between the relevant stakeholders in Spain. Furthermore, no target levels are set (e.g. expected performance results). The link between the statistics results and policy design tends to be weak. Finally, the statistics only address limited segments of the policies' result chains (broadly overlooking the intermediary and final outcomes of policies).

The monitoring framework has been somewhat better established for employment services and ALMPs than in the other policy fields related to social and labour market integration in Spain. Royal Decree 3/2015² and other related regulations have set the co-ordination instruments of the national employment services system, facilitating data exchange between the AACC and the Spanish public employment service (SEPE) and agreements on shared objectives and indicators. In addition to usual indicators on policy take-up, client satisfaction has been observed among the monitoring indicators for employment services and other ALMPs, such as training. As employment data from the social security register (Social Security Treasury under MISSM, TGSS) are linked with jobseeker data in the information technology (IT) system of SEPE (SISPE), the monitoring framework for employment services also has the potential to include outcome indicators. Yet, this potential has not been fully realised for monitoring or evaluation purposes.

It has been more challenging to produce comprehensive national-level statistics and set a monitoring framework regarding the provision of social services in Spain (OECD, 2022^[4]). The statistics on social services entirely funded by the AACC (such as specialised social services) are particularly problematic as these services are not regulated by national laws or agreements setting out obligations for producing the relevant statistics. The integration of data from the AACC IT systems into national-level IT systems (the Information System of Autonomy and Care System for Dependent People [SISAAD] and the Information System of the Users of Social Services [SIUSS]) has been cumbersome due to a lack of common taxonomies; difficulties interpreting and implementing the General Data Protection Regulation; lack of resources to implement data exchanges; and, at times missing unique personal identities (Fernández, Kups and Llena-Nozal, 2022^[6]).³

Although some references to evaluation appeared in Spanish regulations before Law 27/2022 on the Institutionalisation of Public Policy Evaluation, ministries' evaluation activities relative to the labour market and social inclusion have been limited. The reasons for limited evaluation activities are broadly similar to those affecting the implementation of monitoring frameworks, such as challenges in data exchange across governance levels; limited data exchange across national level registers; difficulties in interpreting and implementing data protection regulation to enable M&E activities; modest capacity; skills and experience of the units for data analysis in the ministries to conduct evaluations; and a less developed policy evaluation culture than in some other OECD countries.

The three ministries in charge of policies related to social and labour market integration in Spain do conduct some ex ante and ex post evaluations of policies and legal acts (OECD, 2020^[3]; 2022^[4]), particularly when new policies are introduced. However, these evaluations focus primarily on statistical trends and not on more advanced econometric methods and largely address budgets and take-up numbers (input and output variables), not the policies' ultimate objectives (outcomes). As such, advanced evaluations, like

counterfactual impact evaluations and process evaluations, are not conducted systematically but only occasionally in co-operation with external partners [see, e.g. an advanced impact evaluation of a digital tool for employment counsellors conducted by the OECD in the context of a technical support project by the European Commission's Directorate General for Structural Reform Support (OECD, 2022^[7])].

Regional and local-level practices of M&E social and labour market inclusion

As highlighted in Chapter 2, M&E approaches in social and labour market inclusion vary significantly across AACC. Regarding social inclusion plans, nine AACC foresee a comprehensive evaluation, while four assess each social inclusion programme individually. Within AACC, at a local level, M&E predominantly involves annual reports that collect information on user profiles, programme activities, objectives and performance indicators. Overall, evaluations tend to focus on outputs, with impact evaluations and user experience monitoring largely absent.

The expertise in M&E among municipalities and regions has improved over the past years, largely due to the European Social Fund (ESF), which requires all projects financed through the ESF to adhere to a specific M&E framework that calls for conducting ex ante, intermediary (The National Strategic Evaluation Plan for the ESF 2014-20)⁴ and ex post evaluations of the programmes. This framework also establishes common performance indicators to be monitored, encompassing aspects such as participants, entities, intermediate results and long-term outcomes.⁵

5.1.2. Methodology

This section discusses the M&E activities that MISSM has conducted since the introduction of the new IMV scheme and puts forward actionable guidelines to develop a new concept of a systematic M&E framework for the SIM of the IMV beneficiaries.

National-level M&E introduced with the new minimum income scheme

As practices of linking data across policy fields for M&E purposes were relatively modest when preparations for the IMV started (see the previous section), MISSM had to make considerable efforts to access sufficient data to conduct ex ante evaluations of this new benefit scheme. As a result of these efforts, it was possible to use considerably elaborate datasets for the ex ante evaluations, involving micro-level data from the Tax Agency (Agencia Tributaria, AEAT), the National Statistics Office (Instituto Nacional de Estadística, INE) and the Income and Living Conditions Survey from 2018 regarding País Vasco and Comunidad Foral de Navarra, covering in total 16.3 million households in Spain. The datasets were prepared jointly by the AEAT and INE and were provided to MISSM to ensure reliable data on the target population of the IMV scheme.

The ex ante evaluation of MISSM centred mainly around the first of the two objectives of the IMV scheme, which is income redistribution with a focus on poverty eradication. The evaluation was meant to inform the design of the IMV scheme and, as such, various thresholds of the main design elements (like income and wealth thresholds) were tested to understand the potential take-up numbers, benefit amounts, total financial costs and the changes in poverty rates (extreme poverty, high poverty, moderate poverty, income improvements), i.e. the objective of the new scheme. The datasets used by MISSM included data on household composition, which facilitated designing the new scheme, as well as evaluating the indicators by household types (the number of adults and children in a household).

The second objective of the IMV scheme, which was to address social inclusion and labour market participation, was not included in the ex ante evaluation of the scheme, partly due to the difficulties in designing the specific policy elements related to social and labour market inclusion and the devolved competencies on this matter, but partly also due to limited data available to MISSM. The dataset available

for analysis only enabled the examination of different types of income of the households, including identifying households without any income from labour, i.e. jobless households.

The indicators developed within the ex ante evaluation have been used internally in MISSM to monitor the IMV scheme. Some elements of the policy design, such as the associated work incentives, have changed somewhat since the scheme's introduction (also drawing on the monitoring results), making comparing the current scheme's results and the indicator values from the ex ante evaluation difficult. In addition, a few other variables are monitored to understand whether and how the scheme functions, such as the non-take-up rate, the number of people receiving MIS that were assumed not to meet the eligibility criteria according to the AEAT-INE database used for ex ante evaluations, time duration between application and granting or rejecting the benefit.

MISSM currently publishes the cumulative number of IMV beneficiaries and is considering publishing some of the monitoring indicators of the IMV scheme in the future, creating a better understanding of how the scheme functions and meets its objectives. In addition, some indicators on the IMV scheme are already published by AIReF in its evaluations, looking at somewhat similar indicators as MISSM, such as take-up, coverage and the effects of the outreach campaigns (AIReF, 2022^[8]; 2023^[9]).

MISSM aims to further develop its internal framework of variables to better understand, for example, the issue of non-take-up of the main IMV benefit and the child component to be able to finetune the benefit scheme further in the future. However, no specific indicators have been explored yet to better understand social and labour market inclusion.

A considerable hurdle for MISSM to further develop the monitoring framework for the IMV and social and labour market integration has been the limited access to (linked) administrative data. Although the legislation adopting the IMV scheme also introduced regulation to access all relevant data for ex ante and ex post M&E activities, the data that MISSM accesses currently are similar to those available before the scheme's introduction. These include regular updates on a dataset, including, for example, AEAT data and employment history data on IMV recipients, but no adaptations for the specific purposes of monitoring the IMV scheme or data from SEPE on labour market interventions to monitor labour market integration pathways.

High-level monitoring framework

The OECD defines monitoring as an ongoing function using systematic data collection on specific indicators. This provides key stakeholders indications on progress toward objectives and the use of allocated funds. The key components of a monitoring framework are the predefined indicators, which can be quantitative or qualitative variables to provide simple and reliable instruments to measure achievement, reflect changes caused by a policy or assess performance (OECD, 2002^[10]).

The objective of setting up a monitoring framework for the new SIM of the IMV recipients should be to enable MISSM and other stakeholders to operatively ensure that the resources to support the IMV beneficiaries are available, the steps necessary to integrate this group are made, and the different related interventions are successful. Such a monitoring framework would make it possible to swiftly identify whether there might be some specific challenges in the different elements within the design and implementation of SIM and address these promptly. The monitoring framework needs to be complemented by systematic evaluations to have credible evidence on the integration model's impact, causality and cost-effectiveness and support strategic decisions on policy design and implementation (see the next section for more details).

Ideally, the monitoring framework for the SIM in Spain should be set up in co-operation and consultation with all relevant stakeholders (MISSM, MITES, MIDSA, SEPE, AACC and potentially others). A jointly agreed framework would ensure that all relevant dimensions of the social integration framework would be sufficiently reflected in the monitoring framework; increase the accountability of the stakeholders to

implement the inclusion model; raise awareness of the importance of their roles for the success of the model; and indeed facilitate reacting promptly in case of emerging challenges reflected in the monitoring results. If a jointly agreed monitoring framework would not be feasible, MISSM would still benefit from it by having greater knowledge of the functioning of the SIM and its challenges. Acting on the identified challenges would be somewhat more limited.

A monitoring framework needs to be sufficiently systematic and comprehensive to support key stakeholders with the necessary operational knowledge. This is why using the theory of change in the form of a results chain (as depicted in Figure 5.1) is often used as the backbone of a monitoring framework, including a thorough set of relevant indicators for each link in the chain. It means that the monitoring framework for the SIM should ideally cover these key aspects of providing services and measures for the IMV beneficiaries. However, introducing some of these elements might be challenging at the AACC level, and agreeing on a harmonised and granular enough methodology might take time:

- resources necessary to implement the inclusion model (budget and staff for social, employment and potentially other services relevant to integrating IMV beneficiaries, as well as resources for co-ordination across service providers, their outreach activities, etc.)
- activities necessary to be carried out by the different service providers (promotion, outreach, co-ordination activities, etc.)
- beneficiaries of the interventions of the service providers (e.g. recipients of training or social housing)
- intermediary steps towards social and labour market integration (e.g. improved skills and employability of IMV beneficiaries)
- integration into society and the labour market (e.g. sustainable employment).

Indicators in the monitoring framework should ideally involve quantity and quality aspects in implementing the SIM for a full comprehension of progress in supporting IMV beneficiaries. Indicators addressing quantities generate knowledge on issues like budgets spent on social services or take up of training by IMV beneficiaries. Quality indicators support quality management of the SIM and help make it possible to understand the reasons behind gaps between budgets and expenditures, gaps in take-up rates of integration pathways or high drop-out rates from specific interventions. Monitoring quality aspects of the outcome indicators helps, for example, ensure that the SIM would support IMV beneficiaries in moving to good quality jobs and not push them into low value-added employment. To include quality aspects in the monitoring framework, using survey data in addition to administrative data becomes particularly relevant (see also Section 5.1.4).

As IMV beneficiaries can face very different obstacles to social and labour market integration and need different types of services and measures to overcome these obstacles, the indicators along the results chain should be monitored by the relevant types of support to detect which components might need to be strengthened. To further understand the mechanisms of integration, it might be relevant to study some support measures in further detail (e.g. not only training provision generally but key types of training individually). Furthermore, monitoring the different (sub)interventions by participant subgroups would ensure that policies reach and support the target groups without discrimination or creaming.⁶

Ideally, a narrower set of key performance indicators (or “royal” indicators) should be identified within the overall monitoring system to enable stakeholders to quickly comprehend whether the overall social integration model delivers the main expected results or not. The wider set of detailed monitoring indicators would then serve as a broader background for the key performance indicators to identify the exact implementation challenges and finetune these aspects, if necessary, to improve the overall performance results. Thus, easily accessible (e.g. via dashboards) and understandable key performance indicators are crucial for the strategic management of the social integration pathways, particularly by the higher level officials, and can serve as the main accountability framework for the inclusion model. Ideally, the key

performance indicators should have a focus on measuring final outcomes and have assigned (anticipated) target levels (such as “SMART” targets: specific, measurable, achievable, relevant and time-bound), as well as a framework to address challenges in case of under-achievement and encourage over-achievement. The key performance indicators would have particularly high value-added to drive the performance of implementing the SIM should they be agreed upon between all key stakeholders implementing the SIM.

It should be noted that such agreements would need considerable good will and efforts from all sides in the current set-up in Spain. While some key performance indicators could be developed based on national-level administrative or survey data, some might need to be developed based on AACC data. For this, agreed harmonised methodologies are essential. The evaluations of SIM (discussed in the next section) would complement the key performance indicators within the overall accountability framework with further credible evidence on the model’s performance. However, they would be less timely and frequently available, as conducting evaluations takes more financial resources, skills and time.

A matrix for a high-level monitoring framework for the SIM is proposed in Table 5.1, providing specific examples of possible relevant indicators. The matrix proposes a systematic approach to considering different relevant aspects of monitoring. However, not all dimensions might be needed or feasible in the Spanish case, e.g. due to the governance model and the exact objective of the monitoring framework that still needs to be agreed upon by the stakeholders. When developing the new monitoring framework, MISSM needs to establish indicators for further relevant interventions, sub-interventions and subgroups of the IMV beneficiaries to be included in the framework in consultation with the other stakeholders. Feasibility in terms of data, reliability and timeliness of the indicators should also be considered. Ideally, the exact monitoring framework should be discussed with the different stakeholders while agreeing on the design of SIM. In addition, it might be relevant to increase the coverage of the monitoring framework (in the future) to further policy fields relevant in supporting IMV beneficiaries, such as health, education or housing measures.

Table 5.1. High-level monitoring framework for the social inclusion model of IMV beneficiaries

Examples of possible indicators to monitor social services and ALMPs using the theory of chain approach

| Results chain | Inputs | Activities | Outputs | Intermediary outcomes | Final outcomes |
|--|---|--|--|--|---|
| General definition | Resources for the policy | Actions to convert inputs into outputs | Services, measures and benefits provided | Results of the policy to support the ultimate objectives | Ultimate objectives of the policies |
| Definition for the social inclusion model of IMV beneficiaries | Resources to implement the social inclusion model | Activities of the various stakeholders to organise the delivery of services and measures | Services and measures provided for IMV recipients to help them integrate | Progress in removing IMV beneficiaries’ social and labour market obstacles | Social and labour market integration of IMV beneficiaries |
| Indicators for labour market integration (quantity) | Active labour market policies (ALMPs) expenditures Expenditures on training Number of counsellors | Assessing skill needs in the labour market and organising respective training courses | Number of IMV beneficiaries receiving job counselling Number of IMV recipients receiving training | Increased job search Increased skills | Rate of IMV beneficiaries entering employment Rate of former IMV beneficiaries who re-apply to the IMV after an employment spell |
| Indicators for labour market integration (quality) | Competencies of counsellors | Assessing skill needs on the labour market using credible labour market information | IMV recipients’ satisfaction with training | Increased skills meeting labour market needs | Employment quality (labour income, sustainable employment, career progression) |

| Results chain | Inputs | Activities | Outputs | Intermediary outcomes | Final outcomes |
|--|---|--|---|--|--|
| Indicators for social integration (quantity) | Expenditures on social services Number of case workers | Organising procurements to provide specific social service | Number of IMV recipients benefitting from care services | Decreased time spent on care responsibilities | Reduced social exclusion Increased social relations Increased physical and mental health |
| Indicators for social integration (quality) | Competencies of case workers | Training case workers to identify social integration obstacles | IMV recipients' satisfaction with care services | Higher self-esteem and motivation for social integration | Increased life quality (well-being and life satisfaction) |
| Indicators for co-ordinated service provision (quantity) | Expenditures on co-ordinated and mixed interventions | Co-ordination meetings between social and employment services | Number of joint assessments by social and employment services Number of joint individualised inclusion plans | Combined indicators for the labour market and social integration above | Combined indicators for the labour market and social integration above |
| Indicators for co-ordinated service provision (quality) | Competencies of the multi-professional teams | Quality of co-ordination meetings | Quality of joint individualised inclusion plans | Combined indicators for the labour market and social integration above | Combined indicators for the labour market and social integration above |

Note: Additional policy areas, subinterventions and subgroups might need to be monitored. The exact methodology of indicators will need to depend on data availability. Some indicators will need to depend on survey data (e.g. satisfaction or well-being) or assessing samples of administrative data (e.g. some of the quality indicators).

Source: OECD internal methodology.

High-level evaluation framework

Types of evaluation

As detailed in the previous section, monitoring is essential to ensure that inclusion programmes are progressing as planned; that activities have been implemented and deadlines respected; that participants and staff are satisfied; and that objectives are being met. Such monitoring makes identifying problems, reacting quickly, and providing appropriate solutions easier. In tandem with monitoring, evaluations serve as a systematic review of an ongoing or completed project, programme or policy, encompassing its design, execution and outcomes (OECD, 2002^[10]). They demand a higher investment of time and resources than conventional monitoring indicators but are nonetheless critical to ensuring inclusion programmes' efficacy. Evaluations measure the success of a programme using distinct evaluation criteria.

The OECD evaluation guidelines, widely accepted as standard, were updated in 2019 following a global consultation process. The revised guidelines encompass six distinct criteria: relevance, coherence, effectiveness, efficiency, impact and sustainability (OECD, 2020^[11]). Given these criteria, a fitting evaluation method can be selected. Although numerous evaluation types exist and they can be classified in multiple ways, one proposed categorisation is as follows:

- **Formative evaluations:** Ex ante assessment of whether a programme or intervention is feasible, appropriate, and acceptable before it is fully implemented. It is mostly appropriate to assess the evaluation criteria "relevance".
- **Process evaluation:** Determines whether programme activities have been implemented as intended. Conducted to assess the "coherence" criteria. As part of this evaluation, user experiences could be incorporated to better understand the effectiveness of programme implementation from the perspective of those directly involved.

- **(Intermediate) outcome evaluation:** Measures intermediate programme effects in the target population by assessing the progress in the outcomes or outcome objectives the programme aims to achieve.
- **Impact evaluation:** Assesses programme effectiveness in achieving its ultimate goals.
- **Cost-effectiveness and cost-benefit evaluation:** Examines the programme's outcomes (cost-effectiveness) or impacts (cost-benefit) in relation to the costs of implementing the programme and, if possible, the opportunity costs for beneficiaries (e.g. foregone earnings) as well as indirect costs for non-beneficiaries (e.g. negative externalities).

While all types of evaluations provide valuable insights, impact evaluation – and specifically counterfactual impact evaluation – holds a unique position within the evaluation toolkit. Its importance will be examined in subsequent sections.

Counterfactual impact evaluations and the different methodologies

Determining the impact, or causal effect, of inclusion programmes on participants is a fundamental aspect of evaluating the new SIM. Indeed, understanding the true effects of the different elements of the new model is fundamental to deciding whether and how these elements should be modified and if it is relevant to pursue them and implement the model at a large scale.

Evaluating the causal impact of inclusion programmes poses several methodological challenges. Mainly, it is not sufficient to measure the employment rate (or other outcomes) of beneficiaries and compare it to that of non-beneficiaries. Beneficiaries are often “selected”, meaning they do not have the same characteristics as non-beneficiaries and are therefore not comparable. To isolate the causal effect of inclusion programmes, it would be necessary to be able to compare the outcomes of its beneficiaries with the outcomes that these same individuals would have obtained if they had not participated. Obviously, it is not possible to observe what would have happened to the beneficiaries in the absence of the programme. To overcome this challenge, counterfactual impact evaluations (CIEs) aim to compare the outcomes of programme participants (the treatment group) with those of a set of individuals as similar as possible (the control group (OECD, 2020^[3]; 2020^[12]). The only difference between the treatment and control groups is that the latter did not participate in the programme. The control group, therefore, provides information on “what would have happened to the individuals receiving the intervention if they had not been exposed to it”: the counterfactual case.

Two types of methodologies can be distinguished among CIEs: experimental evaluations, also known as randomised controlled trials (RCTs) and non-experimental (quasi-experimental or observational) evaluations. In an RCT, the treatment and control groups are selected randomly from a given population. If the selection process is truly random, the characteristics of the individuals in the two groups do not differ on average: the groups are therefore statistically equivalent. Thus, comparing the outcomes of these two groups at the end of the programme enables the isolation of the programme's causal effect. While RCTs are an ideal framework for evaluating public policies, they are not always feasible. They require substantial financial and logistical resources; they can raise ethical concerns; and programmes are often designed without following an experimental protocol. However, there is still a need to evaluate such programmes, and to this end, non-experimental methods can be considered. Those methods essentially attempt to mimic the randomisation process described above by constructing a control group as close as possible to the treatment group so that they are statistically equivalent *ex ante*. For more details on CIE, its different methodologies and the framework for CIE at MISSM and MITES in Spain, see OECD (2020^[3]). Concrete examples of the use of these different methodologies in the context of the new SIM can be found in Section 5.2.

Recent experience in evaluating social inclusion programmes at MISSM

Prior to 2020, CIEs were scarcely conducted at MISSM. However, the Resilience and Recovery Plan (RRP) implemented in response to the coronavirus (COVID-19) pandemic, specifically Investment 7 under Component 23, which aims to promote inclusive growth by linking inclusion policies to the IMV, provided MISSM with the opportunity to develop innovative inclusion policies and promote the use of impact evaluations for evidence-based policy making. MISSM became the first public body in Spain to initiate a large-scale pilot programme using RCTs for public policy impact evaluation. The project's primary objective is to foster the development of inclusion itineraries to support the transition of IMV beneficiaries and other vulnerable groups at risk of social exclusion towards social and labour market integration. The evaluation, considered since the design phase of the programmes, aims to provide evidence of the programmes' efficiency and to facilitate mutual learning and informed decision making on whether to expand the most effective programmes (the "best practices") countrywide.

A total of 34 pilot programmes were launched in collaboration with various partner institutions: 16 with AACCC, 4 with municipalities and 14 with third-sector institutions. MISSM and its partners had distinct and well-defined roles in the pilots. MISSM provides continuous support to partners in the design of the intervention and its implementation and is responsible for designing the impact evaluation (i.e. RCT technical elements) and conducting the impact assessment (i.e. data processing and analysis). To ensure flexibility and adaptability to the diverse realities and challenges across the Spanish territory, partner institutions were given some autonomy in the choice and design of the intervention. They determined the activities to be conducted, the target population, the duration of the intervention and the specific geographical location in collaboration with the MISSM. As a result, the interventions encompassed a wide range of content, including social integration, labour market integration, education, housing and digitalisation. Partner institutions were responsible for estimating pilot programme costs and managing their implementation (e.g. organising and contacting participants). Additionally, they were responsible for collecting the necessary data for analysis and sharing it with MISSM.

By employing RCT methodology in the pilot programmes targeting IMV beneficiaries, MISSM established a robust foundation for evidence-based decision making. This experience can serve as a starting point for reflection when constructing the evaluation framework for the new inclusion model and inclusion policies in general.

Although evaluations must be systematically conducted, it is essential to determine the appropriate evaluation types for specific contexts. In the pilot programme experience, RCTs were preconised as the methodology to be deployed, and the partner institutions designed the interventions. Some partners opted for already existing programmes, recognising an opportunity to showcase their effectiveness and potentially expand them across the territory. However, established programmes might benefit from alternative evaluation methods, especially those whose design is not compatible with RCTs. Finding a suitable control group can be challenging, and ethical concerns may arise due to the difficulty in justifying the exclusion of a portion of the population from an established programme solely for evaluation purposes. Conversely, other partners chose to test innovative approaches to address priority issues (e.g. tackling child poverty in Galicia). These interventions often involved multiple components and posed delivery complexities, further challenging the RCT methodology. This underscores the importance of designing evaluations in tandem with intervention design and implementation.

To maintain transparency, potential pilot programme participants were informed about the RCT methodology, their participation in an evaluation, and the possibility of not benefiting from the programme. However, this set-up may have led to unintended behavioural effects, such as the Hawthorne effect, where individuals change their behaviour simply because they know they are part of the treatment group, or the John-Henry effect, where control group individuals alter their behaviour due to feeling excluded from the programme. To increase the likelihood of obtaining necessary data for evaluation, individuals in the control group were often given incentives to complete online questionnaires.

Additionally, RCTs can be expensive and time-consuming, and their results may not always be generalisable. Alternative evaluation methods, such as quasi-experimental designs or observational studies, may be more fitting for evaluating established or complex programmes. However, non-experimental methods remain largely unexplored in the evaluation of inclusion policies in Spain.

In summary, the RCT pilot experience represents a significant first step towards producing reliable evidence to inform policy making in Spain, as outcomes are evaluated rigorously and objectively. The experience showcased a unique example of collaboration across various institutional levels (state, regions municipalities and third sector). It heightened the awareness of various stakeholders regarding the importance of evaluations for evidence-based policy making while also providing training on the design, implementation, and management of impact evaluations. To ensure this experience is not an isolated event, it is vital to establish a systematic evaluation framework for social inclusion policies, taking into account lessons learned in determining when RCTs are the most appropriate evaluation method and how to execute them efficiently. With a comprehensive understanding of past experiences, establishing a systematic evaluation framework for social inclusion policies can be effectively pursued. Additionally, non-experimental CIE methods should be consistently employed, particularly when budgets are constrained, RCT rigour is compromised, or ethical concerns arise.

5.1.3. Necessary human resources

This section discusses the requirements for staff and skills in MISSM to be able to introduce a systematic M&E framework for the SIM, as proposed in the previous section. In addition, this section outlines how the work of the M&E unit in MISSM could be further strengthened by co-operating with external researchers.

Staff requirements

Before the introduction of the IMV scheme, the different units for statistics and analysis in the policy areas of labour market and social security (around 100 staff members across 7 units in MISSM and MITES) were generally not equipped to apply more advanced econometric methods and conduct policy evaluations (OECD, 2020^[3]). About one-half of human resources in these units went on preparing and linking data for statistics and analytical purposes (validating, structuring and cleaning data, running quality controls), and only less than half of staff were actually producing statistics and analysis, essentially being able to cover only the critical needs for statistics.

A considerable challenge for the statistics units within labour and social security fields in Spain has been somewhat outdated IT infrastructure and sometimes missing solutions for data analytics (although the situation varies across registers and units). As such, producing statistics and analysis has required more human resources, and as the number of civil servants in the Spanish public sector has been tightly capped, ministries have struggled to have resources for evaluation activities. Also, the skill composition in the units has reflected these challenges, as while staff in statistics and analysis have been generally highly qualified, the skills were, for example, particularly strong regarding Query Languages and MS Excel but low in terms of software packages for statistics and analysis (such as STATA, SAS, R, etc.). The few attempts over the years to introduce dedicated evaluation units within the relevant ministries (such as to evaluate ESF programmes or ALMPs) have not been successful, partly related to constrained hiring policies, salary attractiveness and opportunities in the public sector.

Aiming at evidence-based policy making regarding the introduction of the new IMV scheme, MISSM set up a new dedicated unit to support the IMV design and implementation with analysis and evaluation activities, the Deputy Directorate for Inclusion Indicators and Objectives (Subdirección General de Objetivos e Indicadores de Inclusión, SGOII) in the General Secretariat for Inclusion (Secretaría General de Objetivos y Políticas de Inclusión y Previsión Social). The SGOII was explicitly mandated to conduct ex ante and ex post impact evaluations of inclusion policies and build such capacity relatively quickly.

The hiring plan for SGOII aimed at creating a unit in which all staff will have training and experience in conducting analytical work, including designing policy experiments and pilots and other types of CIEs. By 2023, SGOII has developed into a strong, motivated team with a good theoretical understanding of impact ex post evaluations and some experience in conducting ex ante evaluations and designing RCTs. The SGOII team has been able to increase their expertise on impact evaluations within the short time since their creation, as well as support AACC, municipalities and third-sector organisations in designing and implementing the pilot programmes targeting IMV beneficiaries (see the previous section), thus also building the analytical capacity among other stakeholders.

The SGOII team might need to increase staff to be able to fully implement the M&E activities of the SIM outlined in this report, particularly as, in addition to the actual analytical activities, tight engagement with the different stakeholders is necessary. In addition, it would be beneficial for SGOII to further build its experience in CIE, including in terms of applying quasi-experimental methodologies. As the design of the IMV scheme and the related SIM would be more homogeneously implemented across Spain in the future, it might not be as feasible to evaluate the different aspects by using RCTs. Thus, the skills to apply other methodologies would become more relevant. Should the SGOII team outsource such evaluations (or e.g. only be conducted by AIReF), it would be important for SGOII to be well familiarised with the different methodologies to drive the evaluation and evidence-based policy-making processes, identify credible evidence and pinpoint gaps in the available evidence.

Regardless of SGOII being a good practice example of swiftly building good analytical capacity and evidence-based policy making, the successful development of the M&E framework for the SIM will depend on the staff capacity of other stakeholders as well. For example, the M&E framework will need good data from other registers (see Section 1.4), such as from the social security registers, AEAT and SEPE, as well as AACC registers. So far, establishing these data exchanges has not been entirely successful, as these also require human resources in addition to establishing appropriate legal and technical solutions. Staff from the different organisations need to discuss which data could be exchanged to meet the objectives of the M&E framework the best. The data protection officers need to find legal solutions for the data exchange to be compatible with data protection regulations. The IT specialists need to establish the most efficient technical solutions for the data exchange. All these discussions would benefit from the staff across the relevant institutions being more aware of why evidence-based policy making is important and how establishing data exchanges could support this process.

Co-operation with external researchers

Tapping into the expertise of external researchers could facilitate SGOII in advancing in developing the M&E framework for the SIM and building its analytical capacity. Co-operating with external researchers on evaluation or outsourcing some evaluation activities entirely would enable SGOII to make more human resources available for these activities and thus tackle the challenges of the public sector hiring policies. In addition, such co-operation would enable SGOII staff to learn about the evaluation design and methodology during the evaluation process and gain the relevant expertise, building their capacity for the future.

The SGOII team has already started co-operating with external researchers with considerable success. SGOII has been able to put in place a bilateral agreement with the Center for Monetary and Financial Studies (CEMFI) to conduct the evaluations of the pilot programmes targeting IMV beneficiaries (see Section 5.1.2) without additional financial burden. Further building a network of experts with whom to co-operate more regularly could foster SGOII's M&E activities, even if the co-operation would only be in terms of discussing different topics related to M&E. Such co-operation partners could be INE, AIReF, the Bank of Spain, researchers in academia, as well as international experts and researchers.

A way to increase researchers' interest in supporting SGOII in evaluating social integration models would be establishing data availability for such evaluations in SGOII and making these data securely available

for researchers. In addition to enabling the commissioning of analytical work to third parties, facilitating access to administrative data allows researchers to engage in activities that can overlap with the evaluation needs of SGOII. This can represent a win-win situation for SGOII and the research community, as the empirical research activity on social integration models can grow, building the evidence base for policy making without additional cost. Generally, this has also been the main channel that made the current co-operation between SGOII and CEMFI possible. However, in the current set-up, evidence will be only generated temporarily and by one organisation. Nevertheless, SGOII has initiated mapping the metadata concerning the data relevant for the ongoing pilots, which will enable sharing of the pilot data in the future with other researchers in addition to CEMFI. In addition, SGOII needs to develop specific, clear and transparent guidelines and processes for external users to request and receive these data in collaboration with the data protection delegates to ensure data exchange practices comply with data protection regulations.

However, it is unlikely that only “free” collaboration with researchers and evidence produced by other public sector institutions (above all, AIReF) will be sufficient to sustainably cover all the needs for evidence on social integration models. In addition, conducting all M&E only internally in SGOII might not be sufficient to ensure objectivity and credibility of the results among other stakeholders. As such, MISSM needs to establish systematic public procurement processes to generate additional evidence on some of the key dimensions of SIM for the IMV beneficiaries and allocate sufficient funding for these activities. This means that MISSM needs a broader strategy for the M&E framework beyond a system of indicators, methodology and data. The strategy also needs to guide the aspects of funding the system and the related procurement processes.

5.1.4. Data needs and sources

M&E efforts, including CIEs, heavily rely on data availability. Identifying and obtaining information on programme participants and accessing information to select a control group among non-participants are necessary steps in the evaluation process. Furthermore, tracking individuals over time and observing their social and labour market integration outcomes is an essential component of effective M&E (see Table 5.2).

Table 5.2. Data needs for the M&E of inclusion programmes in Spain

| Data needs | Description | Example |
|----------------------------------|---|--|
| Treatment status | Data to identify individuals who participate and those who do not participate in the programme to be evaluated | Data to identify who an IMV recipient is for the evaluation of IMV |
| Individual characteristics | Data on socio-demographic characteristics and other individual-level variables to ensure that treatment and control groups are comparable | Data on demographic characteristics, such as gender, age, education, residence, nationality and civil status; household-level characteristics, such as family composition and characteristics of family members; history of (un)employment, benefits, etc. |
| Activities/intermediary outcomes | Data to identify the specific activities individuals undertake during the intervention to be evaluated or as a result of it | Data on the specific ALMPs or social programmes undertaken in the context of the evaluation of inclusion itineraries |
| Outcomes | Data on the final outcomes of interest for both programme participants and non-participants | Data on labour market integration (e.g. employment and its quality) and social inclusion and well-being (e.g. children's education, self-esteem or health outcomes) to evaluate the new inclusion model |

Source: OECD internal methodology.

Administrative data serve as the most relevant data source, encompassing the majority of the information outlined above. This type of data is collected, used, and stored primarily for operational purposes but can also be used for research purposes without incurring significant additional costs (OECD, 2020^[12]). Administrative data offer the advantage of covering nearly all individuals relevant to a given study, thereby

reducing the potential non-response bias present in surveys. Moreover, administrative data can be more accurate than survey data when measuring complex or difficult-to-remember individual characteristics (e.g. the end date of the last unemployment episode, benefit amounts).

In Spain, there is considerable variation in microdata ownership and access across AACC and municipalities, with no comprehensive and unified national system in place for inclusion services and their users (Fernández, Kups and Llana-Nozal, 2022^[6]). However, a substantial amount of administrative microdata on individuals and programmes related to social and labour market inclusion is collected by various institutions for operational purposes and could be used for evaluation purposes.

SIUSS is the closest approximation to a unified national data system for social services. SIUSS is an operational tool used at the municipal level in the daily work of the primary care social services teams. It records data on users, diagnoses, social interventions, activities performed by professionals, reports, appointment management, etc. SIUSS enables homogeneity for statistical analysis across the municipalities that use it through a common classification and taxonomy of services. However, its adoption varies considerably across municipalities due to its voluntary nature (Fernández, Kups and Llana-Nozal, 2022^[6]). Regarding the data needs for M&E presented in Table 5.2, SIUSS mainly provides data on activities and intermediary outcomes related to social services usage, individual characteristics, and, to a lesser extent, treatment status (depending on whether the programme to be evaluated is part of the registry or not).

Another promising development, the Single Social History (Historia Social Única, HSU), is still in its early stages and has only been adopted by a few AACC. HSU aims to compile all information on the social care received by each person, utilising a digital tool accessible to both professionals and beneficiaries. In particular, it will enable the collection of information on programmes, services and benefits available and received by users through public services as well as third-sector and private providers. Additionally, HSU will expand the information available on social services users by tracking not only the services they receive but also their needs in terms of social care. It may also integrate information from other systems, such as health systems. This could provide a more comprehensive view of target individuals' characteristics (e.g. their needs and health status) and their activities related to social care (e.g. all programmes they request and the ones they benefit from).

When considering employment services, the situation appears more promising. The Information System of Public Employment Services (SISPE) serves as a unified national data system. SISPE integrates information related to the management of unemployment benefits and ALMPs implemented by PES at both state and AACC levels. As such, data on jobseekers' characteristics and the various services and programmes they use, as contained in SISPE, are received through interfaces with operational databases of the AACC (i.e. initially inputted by employment counsellors). Notably, the time lag between data collection and availability for use is minimal, with SISPE providing near-live data. SISPE also houses data on exits from unemployment and employment contracts (including occupation, contract type and duration). In relation to the data needs for M&E detailed in Table 5.2, SISPE includes all necessary data for evaluating employment-related interventions. Additionally, if the new SIM specifically targets IMV recipients, SISPE should facilitate the identification of the relevant population, as IMV recipients must register as jobseekers at regional public employment services. Compared to SIUSS, SISPE provides information on individuals' characteristics and activities and intermediary outcomes and measures final outcomes. However, SISPE lacks data on social services received by individuals beyond those overlapping with ALMPs, and it does not include information on household composition or other household members, which could be very relevant for categorising beneficiaries.

The National Strategy for the Prevention and Fight against Poverty and Social Exclusion (2019-23) envisions integrating SIUSS and SISPE to streamline actions, benefits and services from both protection systems in fostering inclusion. This enhancement would consolidate most of the required data in one location, not only simplifying the generation of evidence for social inclusion programmes, such as

through CIEs, but also facilitating effective monitoring of these programmes as tracking progress, identifying areas for improvement, and measuring the effectiveness becomes substantially more efficient.

Additional data sources could also be explored to enhance the evaluation of Spain's new SIM. For example, the National Social Security Institute (INSS) collects extensive data on IMV claimants through registration forms. This dataset comprises numerous unique variables related to social inclusion, such as housing, family composition, disabilities, dependency and victimisation due to gender violence, as well as socio-demographic characteristics and employment status and earnings. Data on employment history from INSS is also a valuable resource that can further enrich the analysis. These data could be shared with other institutions, alleviating the burden of multiple organisations collecting similar information. Tax data from the AEAT may also be relevant, as it collects income, wealth and tax variables, along with other derived information, for all taxpayers in Spain. Consequently, this data source could help in gathering information for constructing a comparable control group and obtaining final outcome variables, such as income and wealth, thereby enabling the estimation of material deprivation and poverty levels.

To complement these administrative data sources, survey data could also be collected to evaluate the different elements of the SIM, including outreach, conditionality, referral and assessment, and others. Survey data would provide valuable information on individuals' behaviours, attitudes, or perceptions that are not typically captured in administrative data. For instance, survey data may be more appropriate when evaluating the impact of a programme on individuals' subjective well-being or job satisfaction. Notably, the pilot programmes targeting IMV beneficiaries do include this type of survey data in their evaluations, demonstrating its relevance and utility in the assessment process.

Building upon these existing information systems and data sources, it is vital to strengthen the data exchange infrastructure and data integration mechanisms (Fernández, Kups and Llana-Nozal, 2022^[6]; OECD, 2020^[3]). This is particularly necessary at subnational levels to provide a comprehensive view of social and labour market services. Ideally, a well-designed M&E framework should use digital solutions to produce and distribute results. Data warehouse and data lake solutions can help effectively prepare data from diverse registers for analytics. Business intelligence tools play a crucial role in this context. They provide capabilities such as setting up pre-defined queries, which allows for an efficient, systematic production of monitoring indicator values. Additionally, these tools offer visualisation functionalities that significantly speed up understanding trends and comparisons across various subgroups. Moreover, creating tailored dashboards within business intelligence tools serves to channel information appropriately according to the specific needs of different user groups. This approach optimises the data delivery process and enhances overall efficiency in monitoring and evaluating social services.

Having explored the general framework for M&E of the new SIM, including its objectives, methodology, necessary resources, and data needs, the next section will discuss the practical aspects of implementation. It serves as a roadmap for the pilot implementation of various components of the new SIM, such as referral and preliminary assessment, as well as exploring the conditionality rules involved.

5.2. Roadmap for pilot implementation of selected components of Spain's new social inclusion model

This section outlines the primary steps to conduct a CIE of two key components outlined in Chapter 6: the assessment and referral protocol and the conditionality within the new SIM. The conditionality aspect is illustrated with the example of compulsory schooling for IMV beneficiaries' children. The section defines the outcomes of interest, the results chain, the method, the necessary data and the action plan for conducting the evaluation.

5.2.1. Referral and preliminary assessment

One of the main challenges for social inclusion in Spain is the absence of efficient referral protocols (see Chapters 3 and 4). Currently, there is no systematic assessment or referral procedure for individuals who receive IMV benefits, which leads to many individuals not benefiting from services and itineraries adapted to their needs. To address this issue, the new inclusion model outlined by the OECD proposes to develop a comprehensive assessment and referral protocol that directs IMV recipients to an assessment and referral unit. There are various options for setting up this unit, including a multidisciplinary team comprising employment experts and social workers. This unit will be responsible for assessing the needs of IMV recipients and determining which services would benefit them most. For example, they could be directed towards a labour itinerary, a social itinerary or a socio-labour itinerary. Generating evidence on whether and how setting up a comprehensive and structured referral and assessment protocol leads to better inclusion pathways is crucial to building a new and well-performing SIM.

The conceptual framework and the outcome variables to be assessed

This section first explains how constructing a conceptual framework can help structure the understanding of the objectives of the new assessment and referral protocol. It then describes the outcome variables of interest for its evaluation.

The results chain

To clarify the evaluation questions, it is essential first to understand the causal logic behind the new model's referral and preliminary assessment protocol: from identifying available resources to defining the results to be achieved. More specifically, it involves determining the inputs, activities, outputs and outcomes of the referral (Gertler et al., 2016^[1]). The “results chain” can be defined as follows:

- **Inputs:** The resources available for the referral. The funds, the human resources related to the staff in charge of the assessment and referral, assessment tools, administrative resources, facilities and materials (computers, paper forms, etc.), etc.
- **Activities:** Actions that transform inputs into outputs, including all aspects of assessment and referral, such as identifying individuals' needs and determining the most suitable ones.
- **Outputs:** Tangible goods and services generated by referral activities, such as the number of individuals participating in the referral, the number approaching advised services, and following recommended itineraries, etc.
- **Net outcomes (impacts):** The effects produced by the referral protocol on the target population after exposure, accounting for the counterfactual outcomes of participants had they not participated. These effects can be measured along different dimensions; those of interest for this evaluation are detailed in the next section.

Inputs, activities, and outputs comprise the implementation phase, which is usually the responsibility of the stakeholders in charge of implementation (AACC). Consequently, CIEs primarily concentrate on identifying the net outcomes. The central question is whether any observed changes can be causally linked to the referral process. By evaluating this causal impact, it becomes possible to determine if the referral protocol objectives within the new inclusion model have been met. The subsequent section outlines the outcome variables that can be used to measure the referral's impact.

Definition of the outcome variables of interest

The primary objective of the assessment and referral protocol within the new inclusion model is to enhance participants' social and labour market integration. As a result, the main outcome variables analysed will naturally relate to social and labour market integration. However, depending on data availability (see

Section 5.1.4), a counterfactual evaluation may offer an opportunity to determine whether the intervention affects other dimensions and help understand the mechanisms at play. In this regard, secondary and intermediate outcome variables may also be investigated.

Suggested main outcome variables related to social and labour market inclusion can be categorised into the following categories:

- **Employment/labour market integration:** This variable indicates whether individuals are employed (self-employed or employed) at a specific point in time or during a specific period (e.g. number of days employed in the six months after exiting unemployment or employment status one year after the initial interview with the referral team). It measures whether participants in the new referral process are more likely to find employment due to their participation (and are thus more likely to be employed) than if they had not participated. Participation in a training programme may also be considered a step towards labour market integration, reflecting an aspect of enhanced employability, and is thus included in this variable.
- **Quality of employment:** Beyond determining whether individuals have increased access to employment, it is essential to assess the quality of the jobs found. Several variables can be measured, such as wages, contract stability (permanent, fixed-term, temporary, etc.), contract duration, individual's job satisfaction, etc.
- **Income and wealth:** This category encompasses wages and other revenue types. Income and wealth variables are commonly used to measure social exclusion, as they help identify whether individuals or households fall below a specific poverty threshold and are at risk of material deprivation. If data allow, disposable income, which accounts for potential debts individuals incur, can also be explored.
- **Housing:** An individual's housing situation is crucial to their broader social integration process. Outcomes in this category should measure the quality of housing, evaluating aspects like living conditions, safety and access to utilities, and housing stability, which considers the frequency of moves or periods of homelessness.
- **Social integration:** This aspect focuses on the path towards complete societal inclusion. It reflects outcomes such as access to social services, availability of social connections, improvement in social skills and participation in community activities.
- **Health:** Variables in this category should encompass both physical and mental health. For example, it may be valuable to track individuals' likelihood of having a physical injury, a common disease, a mental health disease or an alcoholism or drug addiction issue.

Additional secondary outcome variables related to inclusion can be explored if data allow. These may include measures of subjective well-being, such as individuals' confidence and satisfaction levels; children's education, including school enrolment, days of absence and grades, etc.

Furthermore, intermediary outcomes should be explored to understand the mechanisms behind the main outcomes' effects within the context of the new referral procedure. These may include:

- **Information on take-up:** Are participants in the new referral protocol approaching services/itineraries advised more frequently?
- **Job search tools:** Do participants use a broader range of tools/resources while looking for a job?
- **Number of applications submitted and job interviews conducted.**
- **Autonomy:** Are participants less reliant on others to act and make decisions?
- **Degree of recipient dependency on IMV:** How long do they stay in the scheme?
- **Cognitive and non-cognitive skills** targeted by the itineraries (e.g. digital skills).

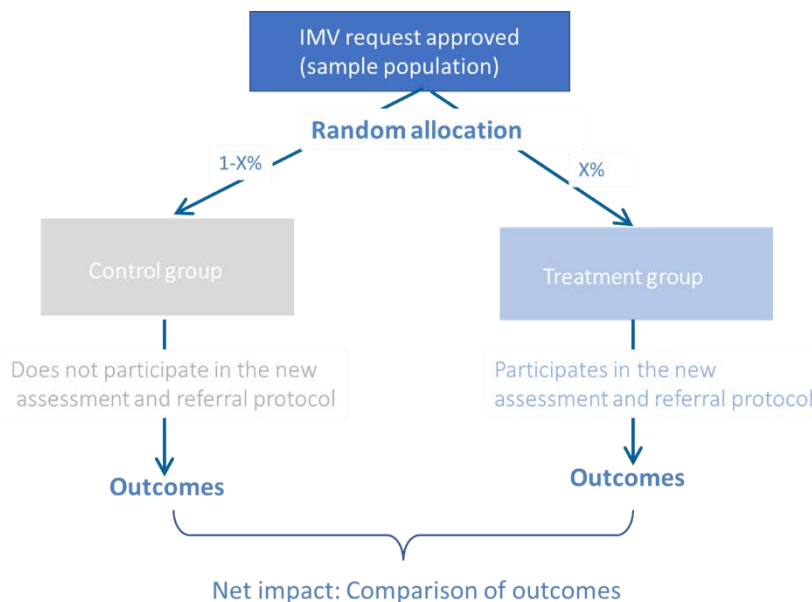
An RCT to evaluate the impact of the new assessment and referral protocol

As discussed in Section 5.1.2 – and in greater detail in OECD (2020^[3]) – RCTs are the most straightforward method for measuring a policy’s causal impact (Duflo, 2007^[13]; Abdul Latif Jameel Poverty Action Lab, n.d.^[14]). They are particularly valuable for piloting policies prior to large-scale implementation.

Piloting the new assessment and referral protocol through an experimental approach appears to be the most accurate choice for evaluation. Setting up the new referral protocol may require substantial resources that could be limited in the short term. Most importantly, human resources needed for assessment and referral units may be scarce. As a result, providing all IMV recipients with the new referral protocol in the short term seems unrealistic. An oversubscription design, where potential participants exceed programme capacity, justifies conducting an RCT since it can be implemented without major ethical concerns and because accurately measuring the referral protocol’s impact is crucial before committing significant resources to it.

The new referral protocol may intervene immediately after the IMV claimant’s request has been accepted. The more straightforward way to proceed is thus to randomly allocate individuals to participate in the referral protocol among the pool of individuals with an accepted IMV request. Figure 5.2 illustrates how the RCT could be conducted. First, a specific proportion of IMV recipients are randomly allocated to the treatment group, while the remaining recipients constitute the control group. The treatment group then participates in the new assessment and referral protocol, while the control group does not. Finally, after the intervention, outcomes for both groups are compared to determine the causal impact of the new protocol. It is important to note that the proportion of individuals assigned to treatment depends on capacity constraints (i.e. the number of referral units that can be set up and the number of individuals they can assist) as well as power calculations that provide information on the necessary sample and treatment size to have sufficient statistical power to detect the expected effects.

Figure 5.2. Diagram of the evaluation of the new assessment and referral protocol through an RCT



Source: OECD internal methodology.

Data requirements to conduct the evaluation

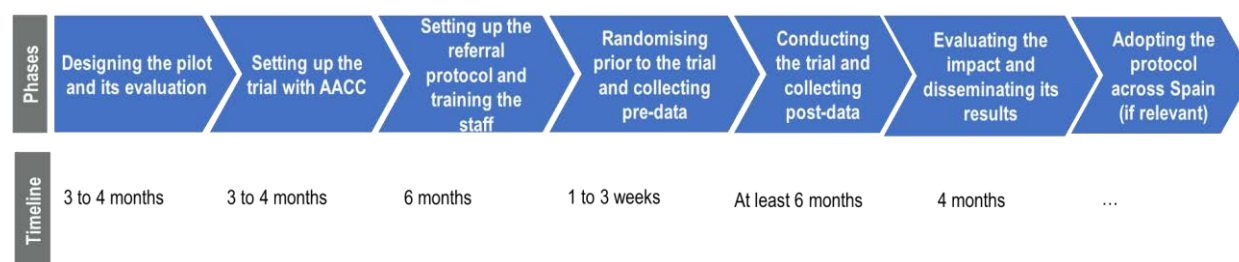
To conduct a rigorous evaluation of the new assessment and referral protocol through an RCT, specific data are crucial. The evaluator should be able to identify individuals in the treatment and control groups, ideally through a unique identifier, and have information on their key characteristics, like level of education, gender and family composition. This information is important mostly to check the balance between the two groups and verify that the random assignment was properly conducted. However, exhaustive data on socio-demographic characteristics or employment history is not as vital in an RCT. The design ensures the groups are statistically equivalent, minimising the need for extensive control variable adjustments. Thus, the primary focus is to collect data on the primary outcomes of interest.

Since IMV recipients are mandated to register at regional PES, SISPE seems to be the most natural primary data source for this evaluation (see Section 5.1.4). Both control and treated individuals can be identified, and their main observable characteristics are recorded (socio-demographic characteristics, unemployment history, etc.). Furthermore, employment-related outcomes are tracked over time for both groups, including intermediary outcomes, such as the use of ALMPs, and final outcomes, such as the likelihood of finding a job and the details of the job found (type of contract, number of hours, sector, etc.). However, these data lack information on outcomes related to social inclusion. Therefore, linking SISPE data with either SIUSS or HSU is needed to assess the effect of the assessment and referral protocol on individuals' social integration. This process should take into consideration the harmonisation and comparability of data. Furthermore, additional data collection could occur at the AACC level as some of them have experience in collecting non-administrative data through surveys for evaluation purposes through recent pilots.

Action plan for pilot implementation

To successfully implement the pilot, several key phases should be followed (see Figure 5.3). The phases could take about 25 months in total. The first phase involves designing the pilot and its evaluation. This initial phase involves the elements discussed in previous sections: determining the results chain and outcomes of interest; selecting the evaluation method while taking into account ethical considerations; identifying the eligible population for the treatment and comparison groups; and determining the appropriate sample size through power calculations. Data needs must also be assessed, and a data collection plan should be developed. Additionally, an evaluation team should be established, and potential risks of the evaluation, such as selective survey attrition and spillover effects, should be assessed.

Figure 5.3. Action plan for the implementation of the new assessment and referral protocol



Source: OECD internal methodology.

Since the evaluation of the new assessment and referral protocol takes place at the level of the AACC, the second implementation phase is to set up the trial with the AACC. This phase involves ensuring that regional stakeholders, such as regional PES and social services, fully understand the process and evaluation needs. Co-ordination is crucial to implementing the referral protocol and collecting data

effectively. After this, the assessment and referral protocol can be set up. This entails establishing all necessary elements for the referral protocol to be functional, such as designating a physical space, hiring or allocating required staff and training staff to conduct the assessment and referral protocol (for instance, familiarising them with the tools to be used to assess the needs and allocate individuals to the different itineraries). It is also essential to raise awareness among staff about the evaluation design and protocol.

Once the assessment and referral protocols are established, randomisation should be conducted before the intervention. The evaluation team can work with MISSM and the AACC to carry out randomisation. Balance checks must be conducted to validate that the treatment and control groups are statistically equivalent before the intervention. It is essential to conduct the randomisation quickly, as newly registered IMV beneficiaries should not wait too long before participating in the referral protocol; otherwise, the efficiency of the protocol could be compromised. At this stage, data collection can take place simultaneously, including baseline observable characteristics and outcome variables prior to the intervention.

The intervention can then be implemented, with individuals in the treatment group participating in the referral protocol while controls do not. The intervention must last long enough for its potential effects to become visible, with six months conventionally being the minimum amount of time required before measuring social and labour market outcomes in this type of intervention. Intermediary outcomes could be tracked before this time horizon. Data collection also takes place during the intervention phase and at the end of it.

After the intervention, the analytical work can begin. The survey or administrative data need to be cleaned and its quality verified. The data can then be analysed descriptively, and subsequently, empirical regressions should be run to retrieve the causal effect of the referral protocol. Robustness checks can be conducted to validate the results. The findings should be compiled in a report and disseminated first internally and then externally. Depending on the results, expanding the assessment and referral protocol to all IMV beneficiaries in Spain could be foreseen.

5.2.2. Conditionality: Compulsory schooling for the children of IMV beneficiaries

In Spain, prior to the implementation of the IMV, regional minimum income schemes (MIS) (Renta Mínima de Inserción) were in place (see Chapter 3). Both the IMV and the MIS share the common goal of providing income to meet the basic needs of individuals or family units. However, the MIS imposes certain requirements, which, despite slight regional variations, generally include residency in the application region, an age range of 25 to 65, prior application for eligible benefits and pensions (e.g. unemployment benefits) and mandatory schooling for dependent minors. Conversely, the national IMV was introduced with no requirements other than the income threshold. This absence of conditionality may have led to unintended consequences, as some individuals might have opted for the IMV over regional schemes to bypass these requirements. Therefore, revisiting some of the conditionality rules is a critical aspect of the new SIM (see Chapter 6). Regions that enforce compulsory schooling for children within their MIS are concerned about the absence of such a requirement for beneficiaries of the IMV. Stakeholders in the education sector have expressed their apprehensions, citing an increase in dropouts among the most vulnerable children. Reinstating this condition could help prevent children from falling into social exclusion that may negatively impact their long-term integration outcomes. Assessing the effects of such a conditionality rule is crucial not only to providing evidence of its importance but also to emphasise the broader significance of schooling for the future integration of vulnerable children in Spain. This section presents a roadmap for evaluating the potential introduction of this conditionality rule.

The conceptual framework and the outcome variables to be assessed

The results chain

The causal logic behind imposing conditionality on IMV beneficiaries to enrol their children in school differs considerably from that of the assessment and referral protocol (see Section 5.2.1). The objectives of these initiatives target different dimensions and individuals and require different resources. For the conditionality rule of enrolling children in school, the “results chain” can be defined as follows:

- **Inputs:** The resources available for putting in place and enforcing the conditionality rule. Mainly the funds and human resources necessary to monitor the fulfilment of the conditionality.
- **Activities:** Enforcing the conditionality rule, thus making sure that the children of IMV beneficiaries are enrolled in school.
- **Outputs:** The number of children of IMV beneficiaries enrolled in school, their attendance rate, their test scores, their future social and labour market integration, etc.
- **Net outcomes (impacts):** The effects produced by the implementation of compulsory schooling on IMV beneficiaries’ children after exposure, accounting for the counterfactual outcomes of IMV participants had they not been subject to the conditionality rule. Net outcomes can be measured along different dimensions; those of interest for the impact evaluation of the conditionality rule are detailed below.

Definition of the outcome variables of interest

The primary objective of mandating schooling for children of IMV recipients under the new inclusion model is to prevent school dropouts among vulnerable children and improve their future social and labour market integration. As a result, the main outcome variables to be analysed will focus on the children of IMV recipients rather than the recipients themselves. They can be divided into short-term and long-term outcomes. Furthermore, spillovers from this policy on integration outcomes of IMV recipients themselves can also be explored.

Regarding the short-term effects of the conditionality rule on children, the suggested main outcome variables can be divided into two categories:

- **Education-related outcome variables:** This first category includes outcome variables such as the probability of IMV beneficiaries’ children enrolling in school, the number of days they attend school, their test scores and their likelihood of failing a subject. These variables measure the policy’s intended first-order effect, which is whether the children of IMV beneficiaries participate in compulsory education and the results they achieve through their participation.
- **Other well-being-related outcome variables, including those for both objective and subjective well-being:**⁷ In addition to assessing whether children have increased their school participation, evaluating other dimensions that affect their well-being and quality of life is essential. Objective well-being variables include children’s physical and mental health status, such as the likelihood of being sick, having a chronic illness or nutritional indicators (overweight, underweight, stunting, etc.). They can also include other dimensions, such as access to necessities, housing quality and stability, and time spent on leisure activities. Subjective well-being could be measured through friends and family relationships, loneliness, feeling safe at school, happiness, motivation, etc. (OECD, 2013_[15])

In the long term, research has shown that additional years of schooling can lead to better integration into the labour market and better pay over the life cycle (Heckman, Lochner and Todd, 2008_[16]; Bhuller, Mogstad and Salvanes, 2017_[17]). Thus, to assess the long-term effects of the conditionality rule, variables such as enrolment and achievement in higher education can be explored, as well as the main outcome

variables related to social and labour market inclusion described in Section 5.2.1 (e.g. employment and income at a given point in adult life).

Furthermore, the enforcement of sending children to school also affects IMV recipients' outcomes directly, which can be analysed through different dimensions. Financial outcome variables, such as investments made in children's education (e.g. textbooks), the impact on IMV recipients' income and wealth, and household task allocation, can be considered. Additionally, the social integration outcomes for parents, resulting from the enrolment of their children in school, may also be explored. For example, they may benefit from interacting with other parents and expanding their social networks.

Method to evaluate policy impact

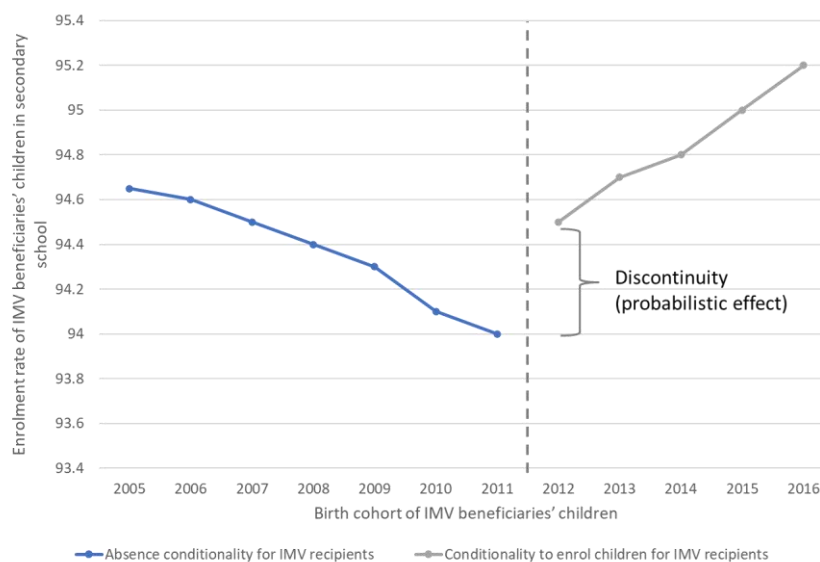
In contrast to the evaluation of the referral and assessment protocol discussed in the previous section, experimental evaluations are not appropriate for assessing the impact of mandating compulsory schooling for children of IMV beneficiaries or any other change in eligibility criteria. The main reason for this includes both ethical and legal concerns. Ethically, randomly assigning individuals to a new conditionality rule raises questions, as it hinders access to a vital resource and a right to a specific subgroup of the population, potentially exacerbating inequalities and causing irreversible negative consequences. Legally, eligibility requirements cannot differ at random, further constraining the feasibility of using experimental evaluations in this context.

The conditionality rule must then apply to all IMV beneficiaries once implemented. This setting requires the use of quasi-experimental methods for evaluation, with regression discontinuity design (RDD) being particularly well-suited. Although the conditionality rule is put in place simultaneously across the Spanish territory, it creates a natural cut-off point in time. Consequently, the degree to which children of IMV recipients are affected by the rule depends on their age (or birth cohort). RDD exploits this cut-off point by comparing individuals just below and above it (Cattaneo, Idrobo and Titiunik, 2019^[18]; Angrist and Rokkanen, 2015^[19]). Individuals on either side of the cut-off are assumed to be sufficiently similar in observed and unobserved characteristics to approximate a random allocation.

The Spanish education system has six years of primary school (from ages 6 to 12) and four years of compulsory secondary education (from ages 13 to 16). This evaluation could focus on compulsory secondary education for simplicity and because children in this age range are more likely to be removed from school for household tasks or work. Additionally, the evaluation should only occur in AACC where existing regional MIS also impose conditionality on enrolling children in school; otherwise, the observed effect could be confounded by strategic decisions to switch from the national to the regional scheme to circumvent this constraint.

IMV was introduced in June 2020 without requiring beneficiaries to enrol their children in school, unlike some existing regional schemes. If the hypothetical date for introducing the new conditionality rule is set to June 2024, all children of IMV beneficiaries of the age to enrol in secondary education from 2020 to 2024⁸ (birth cohorts from 2005 to 2011) would not fall under the new compulsory schooling rule for their parents for at least a year. In contrast, all children of IMV beneficiaries starting secondary school from 2024 onwards (birth cohorts starting from 2012) would be subject to the new rule mandating compulsory schooling. These two groups (treatment and control) can be compared to assess the effect of the conditionality through an RDD (refer to Figure 5.4 for a practical illustration of the methodology). It is important to highlight that the cohorts impacted by the initial lack of a schooling conditionality rule have different levels of exposure. For instance, children from the birth cohort of 2008 were exposed during their entire secondary education, while the 2005 and 2011 cohorts were only exposed for one year. Therefore, it could be beneficial to explore using a continuous treatment variable rather than a binary one to determine the effect of one additional year of exposure to the conditionality.

Figure 5.4. Illustration of a hypothetical RDD setting for evaluating the mandatory schooling conditionality rule



Note: This figure illustrates a hypothetical RDD setting in which the conditionality rule to enrol children in school is imposed on all IMV beneficiaries as of June 2024. This means that children of IMV beneficiaries having the age to start secondary school as of this point in time might be affected; these are children born as of 2012.

Source: OECD internal methodology.

Data requirements to conduct the evaluation

In the context of an RDD or other quasi-experimental or observational evaluations, data requirements are more extensive than in experimental evaluations. In fact, these methodologies attempt to approximate an RCT setting under specific assumptions. To test these assumptions and verify the validity of the methodology, it is crucial to gather a significant amount of data.

The main assumptions in an RDD context are: 1) individuals cannot manipulate the assignment variable; and 2) individuals are equivalent at either side of the cut-off, implying that other variables do not change discontinuously at the cut-off. To assess the plausibility of this assumption, the most common test involves examining whether observed characteristics have identical distributions on either side of the cut-off, meaning no discontinuities should exist in the observables.

Therefore, data on observable characteristics are highly valuable. As detailed in Section 5.1.4, comprehensive socio-demographic and economic characteristics of IMV recipients and their households are collected through registration forms by the INSS. Comparable data could also be obtained from SIUSS, HSU or SISPE sources. Although this information concerns IMV beneficiaries rather than their children directly, it provides ample insight into the children's environment, resources and living conditions, allowing for testing the absence of discontinuity on observable characteristics and thus for the similarity between treatment and control groups.

However, gathering data on the outcome variables of interest presents more challenges. The data described in Section 5.1.4 allow for evaluating social inclusion programmes and policies on their direct target group. In the case of the conditionality rule for enrolling children in school, the main outcomes of interest concern children, not the IMV beneficiaries. In Spain, education policies are highly decentralised. For example, the Spanish Government determines the general characteristics requirements of tests to ensure national standards, but regional governments are responsible for the final design and implementation. Consequently, no national evaluations exist, and education-related data are collected and

stored regionally. To evaluate the conditionality for IMV beneficiaries to enrol their children in school, efforts must be made to link educational (and other) data on children to the data on IMV recipients, which may require co-ordination with the education ministries of regional governments.

Surveys can also be conducted to collect the necessary data, including information related to subjective well-being. However, as this is a long-term evaluation (see the following section), gathering survey data may be costly and challenging. Issues such as attrition, if participants cannot be tracked, or recall bias, if participants provide inaccurate information when reporting past events or experiences, must be considered.

Action plan for the evaluation

Similar to the RCT presented in Section 5.2.1, the action plan's first phase should focus on preparation. This involves designing and planning the CIE and includes key steps, such as determining the theory of change; identifying outcome variables of interest; selecting the appropriate CIE methodology while addressing ethical concerns; assembling an evaluation team; and planning data collection. The latter aspect is vital for assessing the conditionality rule, as potential data sources are numerous and are not yet linked for any other purpose. Therefore, significant co-ordination efforts and investments may be necessary.

Following the evaluation design and identification of data sources, the conditionality rule can be implemented. In contrast to experimental evaluations, there is no need to construct control and treatment groups *ex ante* or train staff on the evaluation methodology. However, to maximise the conditionality rule's potential effects, efforts should focus on making the policy as binding as possible. This entails monitoring the enforcement of the conditionality rule throughout policy implementation.

After the conditionality rule implementation, evaluations can be carried out. Since the primary variables of interest relate to children, a mid-term evaluation may be conducted one or two years after implementation. A second, long-term evaluation could be performed after the first cohort of children is expected to complete secondary school, which would be at least four years post-implementation. Ideally, this evaluation would take place even later, when high-school, higher education, and labour market outcomes can be thoroughly analysed. Reports containing the results from the analysis may then be drafted and disseminated.

5.2.3. Conclusion

This chapter has presented the main steps to conduct a CIE of two key components within the proposed SIM: the assessment and referral protocol and the conditionality of compulsory schooling for IMV beneficiaries' children. The distinct nature of these components requires different evaluation methodologies and timelines.

For the assessment and referral protocol, using an RCT for its evaluation is considered appropriate. This method will accurately measure the protocol's impact before committing significant resources to it, given the considerable human and financial resources needed to establish an assessment and referral unit. It is also worth noting that an extensive amount of data, particularly on the treatment and control groups, is required for such an RCT, underlining the need for data collection mechanisms.

As for the conditionality aspect, given its ethical implications, an RDD is selected as the most appropriate evaluation method. This methodology will allow for an ethically justifiable analysis of the potential impact of introducing compulsory schooling as a condition for receiving IMV benefits. In this case, a comprehensive set of data, not only on IMV recipients but also on their children, is necessary, emphasising again the essential role of data in informing such evaluations.

Despite the different approaches and timelines, these two components should be considered in parallel. Each addresses different aspects of the social inclusion model, and together they can provide a

comprehensive picture of its effectiveness. It is essential to acknowledge that collecting data, especially on social services, is a long-term effort that will require substantial resources and co-ordination. However, such an effort will undoubtedly pay off by supporting the SIM's continuous improvement, thereby enhancing its ability to effectively address social exclusion in Spain.

References

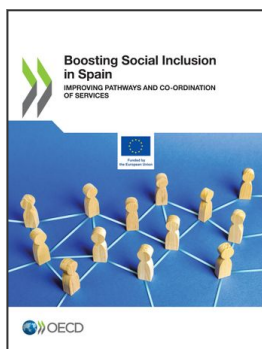
- Abdul Latif Jameel Poverty Action Lab (n.d.), *Introduction to randomised evaluations*, [14]
<https://www.povertyactionlab.org/resource/introduction-randomised-evaluations>.
- AIReF (2023), *2.a Opinión Ingreso Mínimo Vital*, Autoridad Independiente de Responsabilidad Fiscal (AIReF), Madrid. [9]
- AIReF (2022), *Primera Opinión AIReF Ingreso Mínimo Vital*, Autoridad Independiente de Responsabilidad Fiscal (AIReF), Madrid. [8]
- AIReF (2019), *Los Programas de Rentas Mínimas en España*, Autoridad Independiente de Responsabilidad Fiscal (AIReF), Madrid. [2]
- 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. 1/2-1/2, <https://doi.org/10.1080/01621459.2015.1012259>. [19]
- Bhuller, M., M. Mogstad and K. Salvanes (2017), "Life-cycle earnings, education premiums, and internal rates of return", *Journal of Labor Economics*, Vol. 35/4, pp. 993-1/2, <https://doi.org/10.1086/692509>. [17]
- Cattaneo, M., N. Idrobo and R. Titiunik (2019), *A Practical Introduction to Regression Discontinuity Designs*, Cambridge University Press, <https://doi.org/10.1017/9781108684606>. [18]
- Duflo, E. (2007), "Using randomisation in development economics research: A toolkit", *Handbook of Development Economics*, pp. 1/2-3962. [13]
- Fernández, R., S. Kups and A. Llana-Nozal (2022), "Information technologies for social services in Spain: Reform of the national framework for the provision of social services in Spain", *OECD Social, Employment and Migration Working Papers*, No. 277, OECD Publishing, Paris, <https://doi.org/10.1787/f1308a08-en>. [6]
- Gertler, P. et al. (2016), *Impact Evaluation in Practice, Second Edition*, Inter-American Development Bank and World Bank, Washington, DC, <https://doi.org/10.1596/978-1-4648-1/2-4>. [1]
- Heckman, J., L. Lochner and P. Todd (2008), "Earnings functions and rates of return", *Journal of Human Capital*, Vol. 2/1, pp. 1-31, <https://doi.org/10.1086/587037>. [16]
- OECD (2022), *Impact Evaluation of the Digital Tool for Employment Counsellors in Spain: SEND@*, OECD Publishing, Paris, <https://www.oecd.org/els/emp/FinalReport-EvaluationOfSEND.pdf>. [7]
- OECD (2022), *Modernising Social Services in Spain: Designing a New National Framework*, OECD Publishing, Paris, <https://doi.org/10.1787/4add887d-en>. [4]

- OECD (2021), “Building inclusive labour markets: Active labour market policies for the most vulnerable groups”, *OECD Policy Responses to Coronavirus (COVID-19)*, OECD Publishing, Paris, <https://doi.org/10.1787/607662d9-en>. [5]
- OECD (2020), *Better Criteria for Better Evaluation Revised Evaluation Criteria Definitions and Principles for Use*, OECD Publishing: Paris, <http://www.oecd.org/dac/evaluation> (accessed on 2 June 2020). [11]
- OECD (2020), *Impact Evaluation of Labour Market Policies Through the Use of Linked Administrative Data*, https://www.oecd.org/els/emp/Impact_evaluation_of_LMP.pdf. [12]
- OECD (2020), *Impact Evaluations Framework for the Spanish Ministry of Labour and Social Economy and Ministry of Inclusion, Social Security and Migrations*, OECD Publishing, Paris, https://www.oecd.org/els/emp/Impact_Evaluations_Framework.pdf. [3]
- OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264191655-en>. [15]
- OECD (2002), *Glossary of Key Terms in Evaluation and Results Based Management*, OECD Publishing, Paris, <https://www.oecd.org/dac/evaluation/2754804.pdf>. [10]

Notes

1. Ley 27/2022, de 20 de Diciembre, de Institucionalización de la Evaluación de Políticas Públicas en la Administración General del Estado (<https://www.boe.es/eli/es/l/2022/12/20/27>).
2. Further reformed with the new Employment Act adopted on 28 February 2023.
3. The integration of data from the AACC IT systems has been somewhat more successful regarding SISAAD, but very problematic concerning SIUSS. The new national law on social services will likely include an agreement on common taxonomies and more reporting requirements for AACC. However, this draft law is currently on hold.
4. Plan Estratégico Nacional de Evaluación del FSE 2014-2020 (https://www.mites.gob.es/uafse/ficheros/evaluacion/planestrategico-fse/planestrategico_nac_evaluacion.pdf).
5. See EU Regulation No. 1304/2013 (<https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32013R1304>).
6. Creaming occurs when support is provided to those people who integrate into society and the labour market quickly in any case, while those who need the support the most are not able to access it.
7. See OECD’s “Child well-being drivers” at <https://www.oecd.org/els/family/child-well-being/data/drivers/>.
8. Careful consideration is needed for children who were in secondary school in 2020, particularly due to the impact of the COVID-19 pandemic and the resultant school closings in the last quarter of the 2019-20 school year. These unprecedented circumstances may have affected various outcomes, including

students' grades and their subjective well-being. Since this specific cohort could introduce distortions in the RDD analysis, it is advisable to include a robustness check that excludes this group. Doing so will help mitigate any potential biases or inconsistencies in the evaluation.



From:
Boosting Social Inclusion in Spain
Improving Pathways and Co-ordination of Services

Access the complete publication at:
<https://doi.org/10.1787/56b604a0-en>

Please cite this chapter as:

OECD (2023), “An improved monitoring and evaluation framework for social inclusion in Spain”, in *Boosting Social Inclusion in Spain: Improving Pathways and Co-ordination of Services*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/d8c898a6-en>

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. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.