Impact evaluation of labour market and social policies through the use of linked administrative and survey data

Technical report: Impact Evaluation of Vocational Training and Employment Subsidies for the Unemployed in Lithuania

September 2022

Table of contents

Tables

Figures

The report was produced as part of the project “Pilot studies on impact evaluation of labour market and social policies through the use of linked administrative and survey data” which is co-funded by the European Union (European Commission’s Directorate-General for Employment, Social Affairs and Inclusion) (VS-2020-0368).The opinions expressed and arguments employed herein do not necessarily reflect the official views of the OECD member countries or the European Union.

1 Executive summary

This report is undertaken in the framework of a project of the OECD with the European Commission (EC) which aims to raise the quality of the data collected and their use in the evaluation of the outcomes and effectiveness of labour market programmes, so that countries can better evaluate and design policies to benefit their citizens.[1] Within the OECD-EC project, the OECD conducted a counterfactual impact evaluation (CIE) of vocational training and employment subsidies, publishing the results and the consequent policy recommendations in the OECD publication series *Connecting People with Jobs* (OECD, 2022[1]). The current technical report accompanies the report on evaluation results, aiming to build analytical capacity and inform future CIEs conducted by the Lithuanian authorities. This technical report includes a detailed discussion on the monitoring and evaluation framework for active labour market policies (ALMPs) in Lithuania, underlining the importance of CIEs in this framework. It also discusses the process to link and prepare data for analysis, the selection of the appropriate econometric techniques and how these were applied in the CIE conducted by the OECD, and suggests a roadmap to conduct CIEs of ALMPs in Lithuania in the future.

Lithuania has made considerable progress in improving its monitoring and evaluation framework to support evidence-informed policy making. The introduction of the social model has created a strong legal basis for ALMP monitoring and evaluation and has assigned an important role to the Lithuanian Employment Service (LES) for assessing the effectiveness of ALMPs. The monitoring of ALMPs covers the key elements of the results chain (inputs, activities, outputs and outcomes), prioritises the outcome indicators as the key performance indicators, and disseminates the monitoring results publicly. Nevertheless, the monitoring framework has scope for improvements, such as introducing quality elements and adding break-downs for the outcome indicators, and modernising the IT infrastructure to support better data analytics. The Lithuanian authorities are also aiming to strengthen evaluation activities of ALMPs, and for example, several impact evaluations of ALMPs have been conducted over the recent years. To fully support evidence-based policy making, the evaluation activities need to become more systematic and the CIEs need to use experimental methods when feasible, as well as be accompanied by process evaluations and cost-benefit analyses.

The detailed and high-quality data available in the administrative registers offer a rich basis for monitoring and evaluating Lithuania's labour market and social policies. As in many other OECD and EU countries, Lithuania can collect and link information on jobseekers' characteristics, their participation in ALMPs and their employment outcomes by linking public employment service data with social security data. These data sources were used in the OECD's CIE of Lithuania's two main ALMPs, vocational training and employment subsidies. These comprehensive databases facilitated a detailed analysis based on rich information on the personal characteristics of jobseekers (such as their age, education and possible barriers to become employed), their labour market outcomes (notably employment, unemployment, earnings, days worked, occupation) and their participation in various ALMPs. In addition, the available data

[1] "Pilot studies on impact evaluation of labour market and social policies through the use of linked administrative and survey data" which is co-funded by the European Union (European Commission's Directorate General for Employment, Social Affairs and Inclusion) (VS 2020 0368).

make it possible to construct the employment history of jobseekers before becoming unemployed as well as their previous unemployment spells.

The databases used in the OECD (2022[1]) impact evaluation were linked through pseudonymised individual and firm identifiers which masked the identity of the underlying individuals and firms. The correspondence across the databases was generally excellent and linking them proved relatively straightforward. Furthermore, the data used were relatively simple to process, with few records containing contradictory information (for example, overlapping unemployment spells). Statistics pertaining to registered unemployed calculated from the microdata closely match official statistics This suggests that the evaluation results are broadly representative for the registered unemployed.

The choice of methodology used in the OECD (2022[1]) impact evaluation was dictated by both the availability of rich administrative data and the lack of strict ALMP eligibility criteria. In the vocational training and employment subsidy programmes evaluated in Lithuania, participation is not randomly assigned and the conditions for entering them did not contain any strict eligibility conditions. The Lithuanian Law on Employment specifies the target groups of unemployed individuals that are to be the primary beneficiaries of ALMPs and guidelines specify which programmes should be preferentially applied to each of the specific groups. Nevertheless, the law gives LES counsellors the right to exercise discretion in deciding whether to refer an individual to a specific measure and LES counsellors appear to exercise this discretion in practice. Against this background, a dynamic selection on observables econometric strategy was adopted for identifying programme effects.

Going forward, Lithuania could take further steps towards more evidence-informed policy making of its labour market and social policies. These include:

- **Developing the monitoring framework of ALMPs further by including indicators to support quality management**, **targeting employment quality** for jobseekers in addition to employment rates among the outcome variables, and **publishing indicator values by sub-groups** (to ensure that policies reach and support the target groups without discrimination or creaming) **and sub-policies** (to detect which components might need changing). Customer satisfaction surveys could be supported by automatic digital surveys for ALMP participants to collect continuous feedback to feed into quality management.
- **Complementing the monitoring of ALMP outcome indicators by regularly conducting CIEs**. CIEs would measure the effects of ALMP participation after accounting for the counterfactual outcomes that would have been observed had the participants not participated in the programmes.
- **Accompanying CIEs of ALMPs with other types of evaluations for in-depth evidence generation**: formative evaluations for *ex-ante* assessment whether a new ALMP design is feasible, appropriate, and acceptable before it is fully implemented; process evaluations to determine whether activities to provide an ALMP have been implemented as intended; intermediate outcome evaluations to assess the progress in the outcomes or outcome objectives that the ALMP is to achieve; and cost-effectiveness and cost-benefit evaluations to demonstrate whether the funding invested in ALMPs could generate benefits for the society exceeding the investments. Particularly the cost-benefit analyses would help the LES to makes its business case and attract funding for ALMPs. More systematic CIEs of ALMPs in the future will help the LES also conduct systematic cost-benefit analyses.
- **Enriching CIEs with an even wider set of outcomes**. This could entail incorporating information which would include additional aspects such as hours worked, as well as social and other transfers. In addition to allowing for a more nuanced interpretation of programmes' effects, such additional data could facilitate detailed cost-benefit analyses that can take into account the wider effects of participating in the programmes, such as the possibly decreased expenditures on social benefits arising from increased employment rates.

- **Adopting a comprehensive impact evaluation framework to systematically design and implement CIEs**. This would entail understanding the needs for CIEs, determining which programme or intervention to evaluate and for what purpose. This can then inform subsequent decisions on which outcomes to evaluate and how the evaluation is conducted. It can also **allow for experimental approaches** that can be adopted only in conjunction with a programme's implementation, such as by randomly assigning jobseekers to be eligible for a specific intervention (randomised controlled trials).
- **Modernising the IT infrastructure in the LES** to support data analytics, including internal monitoring and evaluation activities, and sharing data for research purposes. As the LES is planning to fully overhaul its current operational IT system, it has a good opportunity to develop good data analytics solutions along the new architecture (including a Data Warehouse, Lake or a Hub solution, as well as a fully functional Business Intelligence tool).
- **Communicating the results of the ALMP evaluations more effectively** in order to secure support and the necessary funding. The latter finding is particularly important in light of the findings of the OECD impact evaluation, which finds that both ALMPs analysed have positive effects on labour market outcomes and make a persuasive case for increasing expenditures on these programmes.

2 Strengthening the monitoring and evaluation framework

Fully-fledged evidence-informed policymaking needs to be developed in the Lithuanian system of active labour market policies (ALMPs) to ensure that policies that are effective in supporting jobseekers and employers achieve sustainable funding. Having credible evidence on the effectiveness of ALMPs and the Lithuania's Employment Service (LES) would help the Ministry of Social Security and Labour and the LES communicate this evidence to the public and policy makers and attract the resources needed to provide ALMPs. Evidence-informed policy making needs to be systematic and involve the whole cycle of designing, monitoring and evaluation frameworks, generating knowledge, disseminating knowledge, adjusting policies based on evidence, as well as evaluating the knowledge generation process itself and adjusting the monitoring and evaluation framework accordingly. Knowledge generation needs to involve a monitoring framework to enable agile overviews on the progress of policy implementation, as well as evaluation activities to generate in-depth understanding of policies and identify causal relationships. Evaluations need to include ex-ante evaluations in designing policies, as well as ex-post process and impact evaluations to understand what works, for whom and how, accompanied by cost-benefit analyses to understand whether the investments in ALMPs bring sufficient value added for the society. Credibly evaluating the impact of policies and assessing their cost-effectiveness allows identifying the need to adapt or terminate inefficient policies and boost the efficient ones. Process evaluations help to design more efficient policy implementation practices. Generating evidence and designing policies based on evidence is not important only regarding specific labour market services (such as the LES counselling services) or measures (training, employment incentives), but also across the tools, processes and approaches that the LES uses.

2.1. Monitoring and evaluation frameworks to support evidence-informed policy-making

Over the years, governments have started to increasingly recognise the importance of knowledge and evidence to inform policy. Thus, monitoring and evaluation frameworks as systematic tools to generate knowledge have gained importance. There are multiple drivers of this increasing demand, including budgetary commitments to increase cost-effectiveness, monitoring and evaluation requirements as part of specific funding arrangements (e.g. European Structural and Cohesion Funds), avoiding erosion of trust in public institutions, and new and complex policy objective and challenges.

Monitoring and evaluation are two pillars of generating evidence that should not substitute each other, but be complementary. Monitoring enables policy designers and implementers to assess evolutions in policies over time, supporting performance and quality management, as well as feeding into operational decisions in addition to strategic management. Evaluation activities look into questions such as whether policies achieve their objectives, whether there are causal links between policies and results or whether the benefits of a policy outweigh its costs. As such, evaluation supports above all decision making on policy design and organisation of implementation. In addition to systematic knowledge generation via the monitoring and evaluation framework, evidence-informed policy-making needs to take into account other, including *ad hoc* information, such as one-off in-depth analyses of policies or evidence generated by external organisations, academia or related evidence in other countries.

2.2. Monitoring active labour market policies needs to be systematic and comprehensive

The OECD defines monitoring as a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing policy with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds. As such, the key components of a monitoring framework are the pre-defined indicators, which can be quantitative or qualitative variables to provide simple and reliable means to measure achievement, reflect changes caused by the policy or assess performance (OECD, 2002[2]).

A monitoring framework needs to be systematic and comprehensive to fully support the management level in the PES and ministry. A common approach is to use the theory of change to achieve a systematic and comprehensive set of monitoring indicators. The theory of change (also called intervention logic) is a description of how a policy is supposed to deliver the desired results.

The most important elements of the theory of change can be presented as a results chain to define the sets of relevant indicators, monitoring the progress of a policy by indicators for inputs, activities, outputs, intermediary outcomes and final outcomes (Figure 2.1, also relevant for evaluation activities, see also Section 5.1.1). A similar framework is for example applied also by the European Social Fund to monitor the progress of policies funded by these resources, dividing final outcome indicators additionally to immediate result indicators (one month after participation in a measure) and long-term result indicators (six months after participating in a measure) (European Commission, 2018[3]).

Figure 2.1. Using a results chain to design a monitoring framework

Results chain	Inputs	Activities	Outputs	Intermediary outcomes	Final outcomes
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 policy
Example on training: measuring quantity	Budget spent on providing training, working hours to organise training	Conducting procurement, selecting providers, referring jobseekers	Number/share of jobseekers referred, participating, completing	Number/share of participants with higher skills, changing job search behaviour	Number/share of participants in employment
Example on training: measuring quality	Competencies of staff organising training	Quality of assessing jobseeker needs, individual action plans	Jobseeker satisfaction with training/providers	Employer satisfaction with candidates to their vacancies	Income, job security and career progression of participants

Source: OECD compilation.

The current monitoring framework of ALMPs in Lithuania largely covers the full results chain. A decree by the Minister of Social Security and Labour (Lietuvos Respublikos socialinės apsaugos ir darbo ministerija, 2017[4]) and an order by the Director of the LES (Užimtumo tarnyba prie Lietuvos Respublikos socialinės apsaugos ir darbo ministerijos, 2020[5]) set the processes and methodology for the monitoring framework (labour market monitoring to be performed by the LES). This methodology focuses on output indicators (registered jobseekers, registered vacancies, ALMPs provided), quality of outputs/activities (client satisfaction) and outcome indicators.[2] The monitoring framework rightfully gives the prominent role for the outcome indicators, which observe employment rates and rates of registered unemployment of participants in training programmes (in total seven different policies) in intervals up to two years after participation, and job maintenance rates of employment incentives (three different policies) up to four years of creating the jobs to assess the LES service provision. Furthermore, the monitoring framework underlines the importance of analysing the trends in the indicator values and comparisons between municipalities to assess progress in ALMP provision, developments in the labour market situation and forecast labour market indicators. Input indicators regarding financial aspects are monitored separately, via budget reports.

Indicators for both quantity and quality need to be monitored for a full comprehension of progress of providing an ALMP. Indicators addressing quantities generate knowledge on issues like budgets spent on ALMPs or take-up of ALMPs. Quality indicators support quality management in the PES and help to understand the reasons behind gaps between budgets and expenditures, gaps in ALMP take-up or high drop-outs from ALMPs. Furthermore, also the PES outcome indicators should aim to cover aspects of quality to ensure that its services and measures fully support a good functioning of the labour market and do not have perverse effects (e.g. pushing jobseekers to low-quality and low value-added employment).

The quality aspects are not strongly present in the current monitoring framework of ALMPs in Lithuania, represented only by surveys for jobseeker and employer satisfaction with the LES. These surveys could be for example supported by automatic digital surveys for ALMP participants to collect continuous feedback to feed into quality management. The outcome indicators regarding employment rates after ALMP participation could be complemented by indicators for job quality, for example building on the aspects studied in the OECD impact evaluation – wages, income and career progression.

In addition, the indicators along the results chain should be monitored by relevant sub-policies (to detect which components might need changing) and participant sub-groups (to ensure that policies reach and support the target groups without discrimination or creaming). The LES produces statistics currently by key characteristics, such as age, gender, working capacity or municipality. However, the LES outcome indicators do not have currently any break-downs by sub-policies (such as by tripartite and bipartite agreements in training provision or by training fields) or sub-groups (such as skill level or unemployment duration) available.

While a comprehensive set of indicators is needed to monitor the progress of ALMPs above all for operational reasons, a narrower set of (key) performance indicators is crucial for strategic management, performance management and the accountability framework. The key performance indicators should have a focus on measuring final outcomes, and have assigned target levels (ideally SMART targets: Specific, Measurable, Achievable, Relevant, and Time-Bound), as well as a framework of incentives to encourage achieving the targets and ensure accountability. Furthermore, the key performance indicators need to be well communicated and accepted throughout the different levels of PES to effectively help the PES achieve its objectives. The Lithuanian key performance indicators are disseminated also publicly (the methodology via the relevant regulations and the results on the LES website (Užimtumo tarnyba, 2022[6])) contributing to a binding accountability framework.

The focus of key performance indicators on outcome indicators is relevant to steer the PES towards effective ALMP provision, supporting jobseekers get good jobs and employers find the skilled labour they

[2] The outcome indicators and client satisfaction indicators are somewhat misleadingly referred to as *evaluation* of the *effectiveness* of ALMP provision and implementation in the regulations, while the methodology does not describe evaluation methodology, but a more simple calculation of monitoring indicators. The described methodology enables the authorities to have some indication of service quality (client satisfaction indicators) and gross effects (share of ALMP participants in employment or in unemployment register after participation), but not net effects and causal relationships between ALMP participation and labour market outcomes. The latter effects can be only established via CIEs, see the discussion towards the end of this section as well as the next chapters, particularly Chapter 5.

need. Historically, PES across countries have focused more on input, activity and output indicators, putting more focus on managing well their assigned budgets than supporting the labour market. Over the recent years, the narrative of effectiveness and making the business case for PES has gained more importance, in addition to prioritisation of outcome indicators in the monitoring framework of the European Social Fund. Thus, the key performance indicators involve increasingly outcome indicators. Also the key performance indicators in the monitoring framework in Lithuania focus rightfully on the outcome indicators.

Nevertheless, the commonly used outcome indicators (such as a share of ALMP participants in employment after a certain period) do not fully reveal the causal relationship between an ALMP and the labour market outcomes of an ALMP participant, or the effects of ALMP provision on the labour market more generally. The causal relationships between ALMP provision and labour market outcomes can be identified using counterfactual impact evaluations (CIEs, as described in detail in the next chapters). In general, outcome indicators using estimated counterfactual effects are simply not easy to include in the monitoring frameworks, where the outcome indicators need to be monitored regularly (annually, quarterly, monthly or even daily) and across PES activities and ALMPs. Nevertheless, with increasingly better access to data for PES and advancing digitalisation, automatic impact evaluations are becoming increasingly more possible and could be part of PES key performance indicators in the future. For example, the German PES has been using a semi-automatic tool called TrEffeR for CIEs already since 2008, and the Estonian PES adopted a similar fully automatic tool just recently (OECD, 2022[7]). Both of these PES monitor the results of CIEs via their digital tools, although more for operational purposes, and have not (yet) included related indicators among their key performance indicators. Still, these developments indicate how evaluation can have more potential to support (or even blend with) monitoring and strengthen PES accountability framework.

An effective and efficient monitoring framework needs to be supported with high-quality and comprehensive data (similarly to evaluation activities, see discussion in Chapter 3) and digital solutions to generate and disseminate the results. Data warehouse and data lake solutions enable using data (potentially across registers) that are well prepared for data analytics. Business Intelligence tools enable set up pre-defined queries for an efficient production of monitoring indicator values systematically. In addition, Business Intelligence tools enable visualisation of indicator values, facilitating immensely a quick comprehension of trends and comparisons across sub-groups. Furthermore, creating dashboards in the Business Intelligence tools facilitates channelling information by user groups according to their needs.

However, the current IT infrastructure of the LES does not support monitoring activities (or data analytics more generally). The LES does not use Data Warehouse or similar solutions nor Business Intelligence tools to produce statistics, and queries for these purposes are implemented directly in the operational IT system. The LES has adopted a Business Intelligence tool Tableau only for disseminating general statistics publicly (currently accompanied by visualisation for a small part of statistics), while the key outcome indicators are continuously disseminated externally from Tableau. As the LES is planning to fully exchange its current operational IT system, it has a good opportunity to envisage good data analytics solutions along the new architecture.

2.3. Evaluating active labour market policies generates credible in-depth evidence

Evaluations are a systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results (OECD, 2002[2]). As evaluations are more time-consuming and costly to implement than the usual monitoring indicators, evaluation results are not generally included in the monitoring frameworks, but support monitoring activities. Nevertheless, systematic evaluations are needed across ALMPs, as well as PES approaches and tools to ensure that these support the labour market in the best possible way.

Evaluations assess the success of a programme or policy based on different *evaluation criteria*. The OECD evaluation criteria are commonly accepted as standard guidelines and were updated in 2019 following a global consultation process. The revised guidelines describe six different criteria: Relevance, Coherence, Effectiveness, Efficiency, Impact, and Sustainability (OECD, 2020[8]). The evaluation criteria are typically the basis to define (more detailed) *evaluation questions* that should be answered by the evaluation.

Given the criteria that the evaluation seeks to address, the suitable *evaluation type* can be chosen. While there are several types of evaluations and different ways of classifying them, one breakdown is as follows:

- **Formative evaluation:** *Ex-ante* assessment whether an ALMP is feasible, appropriate, and acceptable before it is fully implemented. Mostly appropriate to assess the evaluation criteria "Relevance".

  Example for the LES: The evaluation could look at whether a recommender tool on job search advice to be used by employment counsellors is likely to be needed by the counsellors and jobseekers, as well as whether its recommendations are likely to be understood and used.

- **Process evaluation**: Determines whether activities to provide an ALMP have been implemented as intended. Conducted to assess the "Coherence" criteria.

  At the moment, process evaluations are not carried out in the LES and need to be initiated to support the LES in its progress to transform into a modern and efficient provider of ALMPs.

  Example for the LES: The evaluation could look at how the recommender tool on job search advice to be used by employment counsellors is used by the counsellors, and whether the tool is used by the counsellors as intended.

- **(Intermediate) outcome evaluation:** Measures intermediate ALMP effects in the target population by assessing the progress in the outcomes or outcome objectives that the ALMP is to achieve.

  Example for the LES: The evaluation could look at whether the use of the recommender tool has resulted in those who have been counselled using the profiling tool change their behaviour, e.g. changes in job-search activity or enrolling to training programmes.

- **Counterfactual impact evaluation:** Assesses the effectiveness of an ALMP in achieving its ultimate objectives and takes into account counterfactual outcomes. Evaluating public policies and programmes is instrumental for evidence-based policy making. Sound evidence on what works and for whom helps governments to achieve strategic objectives and spending efficiency. Impact evaluations are one of the central parts of an evidence-based policy cycle. They serve as a foundation for greater accountability, innovation, and learning. As a focal type of evaluations, many resources are made available for countries to build the capacity to conduct CIEs of ALMPs: OECD (2020[9]), (2020[10]). European Commission (2019[11]), (2020[12]), (2020[13]).

  While Lithuania monitors employment rates of ALMP participants in its monitoring framework, these do not provide full information on causal relationships between ALMP participation and labour market outcomes. As such, the LES needs to conduct CIEs of ALMPs systematically to support its knowledge generation within the monitoring framework. The impact evaluation conducted by OECD (2022[1]) and particularly the current report highlighting the technical aspects of conducting CIEs can be used for systematic CIEs of ALMPs in the future.

  Example for the LES: The evaluation could look at whether the jobseekers who have been counselled using the recommender tool become employed, earn higher wages, achieve sustainable employment, achieve better occupational match, etc. (after they have potentially changed their job search behaviour and/or participated in training programmes).

  The next chapters of this report will discuss in detail conducting CIEs using non-experimental data and applying a specific quasi-experimental method (propensity score matching using in conjunction with a difference-in-difference estimator). Nevertheless, the gold standard of impact evaluations is via randomised controlled trials (RCTs, see also Section 5.2.1). Hence, the Lithuanian authorities should introduce RCTs in the evaluation framework where this could be feasible, such as when introducing new ALMPs, new digital tools or for example making changes in jobseeker counselling content or frequency.

Example for the LES: In the example above, the jobseekers to be counselled with the recommender tool and those who will be counselled without the tool could be assigned randomly (e.g. by their client ID number). Alternatively, randomisation could be implemented by the counsellor level or local office level. In case of a well-designed and implemented RCT, evaluating the effects of the recommender tool will be more credible, as well as potentially methodologically simpler.

- **Cost-effectiveness and cost-benefit evaluation**: Examines the outcomes of an ALMP (cost-effectiveness) or impacts (cost-benefit) in relation to the costs of implementing the ALMP and, if possible, the opportunity costs for participants (e.g. foregone earnings) as well as indirect costs on non-participants (e.g. negative externalities).

  The LES does not currently conduct cost-benefit analyses. Yet, these types of evaluations could help the LES clearly demonstrate that the funding invested in ALMPs could generate benefits for the society exceeding the investments. These evaluations would help the LES to makes its business case and attract funding for ALMPs. Conducting Cost-benefit analyses can be conducted relatively easily once a CIE has been conducted for an ALMP. As a first step, the LES could complement the CIE conducted by the OECD with a cost-benefit analysis regarding the cost-effectiveness of vocational training and employment subsidies for unemployed. More systematic CIEs of ALMPs in the future will help the LES also conduct systematic cost-benefit analyses.

  Example for the LES: The evaluation could look at whether the benefits from higher employment rates and wages of jobseekers who were counselled using the recommender tool are higher than the cost of developing and implementing the recommender tool in employment offices.

3 The rich set of data sources used could be further augmented

Evaluating the effectiveness of ALMPs requires rich data with detailed information on jobseekers characteristics, their participation in ALMPs and their employment outcomes. The detailed and high-quality data available in the administrative registers provide a rich basis for monitoring and evaluating Lithuania's labour market and social policies. The chapter first discusses the data used in the OECD (2022[1]) impact evaluation of two of Lithuania's main ALMPs, employment subsidies and vocational training. It then discusses some of the limitations and provides suggestions for databases that could be included in future impact evaluations.

3.1. Several data sources were used in the impact evaluation

The data used by the OECD (2022[1]) to conduct the impact evaluation come from several sources and span the period from January 2014 to December 2020 (for an overview, see Table 3.1).[3] Unique individual identifiers allow the data to be combined, providing a rich understanding of individuals' participation in ALMPs, their background characteristics – both from the LES registry - and their labour market outcomes and wages from a number of different sources. This is complemented with additional employment data, covering also those individuals who were not in employment, as well as with firm-level data containing information on the attributes of firms where individuals became employed.

The resulting database contains detailed information on the 947 185 unique individuals who were registered as unemployed at any point during the 2014-20 period. These individuals experienced 2.1 million distinct unemployment spells in total. It also contains detailed information on the 79 700 entries for employment subsidy programme participation and 93 800 entries into vocational training. Individuals entered into vocational training and/or employment subsidies in 7.5% of unemployment spells during this period. The data generally span the period from January 2014 to December 2020. Unique individual identifiers allow the data from the different sources to be combined.

The precise years chosen spanned by the data were dictated largely by data availability. Most notably, data on occupational codes are only available beginning in 2014. Because the question of occupational mobility was an important question to be examined in the analysis, this year was chosen as the cut-off date for the data to begin. This choice nevertheless enabled a suitable range of years available, ensuring seven full calendar years of data.

[3] Information on the variables contained in each of the databases is provided in Annex Table 1, Annex Table 2 and Annex Table 3.

Table 3.1. A rich set of data sources were used in the evaluation undertaken by the OECD

Data source	Information available	Periodicity	Sample	Coverage
Lithuanian Employment Service	Detailed background characteristics of registered unemployed, participation in ALMPs and unemployment benefits	Start and end dates of unemployment spells, participation in ALMPs and unemployment benefit receipt	Registered unemployed	2014-2020
Board of the State Social Insurance Fund under the Ministry of Social Security and Labour (SODRA)	Employment outcomes and earnings	Start and end dates of employment spells	Individuals who were unemployed at some point during 2014-2020	2014-2020
Board of the State Social Insurance Fund under the Ministry of Social Security and Labour (SODRA)	Employment outcomes and earnings	Monthly	Individuals who were never unemployed at some point during 2014-2020	2018-2020
State Enterprise Centre of Registers (Register of Legal Entities)	Business registry data	Changes as reported by legal entities	Registered legal entities	2018-2020

Note: The table provides an overview of the data used in OECD (2022[1]). Although the business registry data contains information also on sole proprietorships, the absence of unique firm identifiers in some of the other data means that they cannot be consistently included in the statistics.

The data provided by the LES contained information on several ALMPs offered during the 2014-20 period (Table 3.2). Panels A and B compare the participant numbers of programmes included in the microdata with those from the EC database; for completeness, Panel C contains additional programmes not included in the microdata. In addition to the employment subsidy and vocational training programmes which were analysed in the OECD (2022[1]) impact evaluation, the data provided contained information on several additional programmes: internships, recognition of competencies, and apprenticeships. However, in consultation with the LES, the decision was made not to include these programmes in the CIE. In contrast to the employment subsidy and vocational training programmes, which each had at over five thousand participants in each calendar year during the 2014-20 period, none of these programmes had more than four thousand participants in any of the years during this period. The largest of these latter programmes – supporting the acquisition of work skills – had as few as 800 participants in a given year. Including the smaller programmes in the analysis carried a greater potential for the analysis being beset by a lack of statistical power – having insufficient observations to make meaningful inferences. This would have increased the likelihood that, for example, the results would not find statistically significant effects for a given programme even if the programme was in fact effective. Nevertheless, these programmes could potentially be evaluated in the future bearing in mind these caveats.

Table 3.2. Participation in ALMPs in Lithuania by programme type and entry year

Comparison of statistics from microdata and EC database (thousands of individuals)

Type of ALMP	2014	2015	2016	2017	2018	2019	2020	Total
Panel A. Calculations from microdata								
Employment subsidies	20.3	14.1	7.1	9.1	10.7	7.2	11.2	79.7
Vocational training	5.8	19.1	17.2	21.4	16.6	8.2	5.4	93.8
Supporting the acquisition of work skills	2.8	3.5	3.2	3.1	2.1	1.7	0.8	17.0
Internship	0.0	0.0	0.0	0.0	0.8	0.7	0.4	1.9
Recognition of competencies - Nonformal	0.0	0.0	0.0	0.0	0.6	0.5	0.3	1.5
Apprenticeship	0.0	0.0	0.0	0.0	0.4	0.3	0.2	0.9
Total, Panel A	28.9	36.7	27.5	33.6	31.2	18.6	18.3	194.8
Panel B. Statistics for comparable programmes from EC database								
Employment subsidies	20.2	14.1	7.1	9.1	10.6	7.1	11.0	79.3

Vocational training of the registered unemployed and employees facing redundancy	5.8	19.1	17.2	21.4	16.6	8.3	5.4	93.9
Promotion of work skills development	2.8	3.5	3.2	3.1	2.1	1.7	0.8	17.1
Internship	0.0	0.0	0.0	0.0	0.8	0.7	0.4	1.9
Recognition of competences acquired by non-formal and informal learning	0.0	0.0	0.0	0.0	0.6	0.5	0.3	1.5
Employment under an apprenticeship employment contract	0.0	0.0	0.0	0.0	0.4	0.3	0.2	0.9
Total, Panel B	28.9	36.7	27.5	33.6	31.1	18.6	18.1	194.5
Panel C. Memorandum item: Statistics for other ALMPs from EC database								
Employment incentives	27.3	22.3	15.4	18.1	18.3	14.4	284.6	400.4
Supported employment and rehabilitation	0.6	0.9	1.1	1.1	0.9	0.5	0.4	5.5
Direct job creation	21.9	21.9	19.3	5.3	0.0	0.0	0.0	68.3
Support for mobility	0.0	0.0	0.0	1.6	7.8	5.1	2.9	17.4
Support of social enterprises	2.6	3.1	3.4	4.2	3.9	4.0	2.6	23.9
Local Employment Initiative projects	1.0	1.2	1.4	1.2	1.3	1.3	0.8	8.2
Job rotation	0.5	0.3	0.1	0.0	0.0	0.0	0.0	0.9
Subsidised employment of the disabled	0.4	0.3	0.3	0.3	0.4	0.3	0.3	2.3
Vocational (occupational) rehabilitation	0.3	0.5	0.8	0.8	0.5	0.2	0.0	3.2
Public works	21.9	21.9	19.3	5.3	0.0	0.0	0.0	68.3
Subsidy for individual activities under a business license	5.8	5.3	5.2	1.4	0.0	0.0	0.0	17.8
Total, Panel C	82.3	77.7	66.3	39.4	33.1	25.8	291.6	616.2

Note: Statistics are based on the calendar year individuals entered into the programme.
Sources: OECD calculations based on data from the Lithuanian Employment Service (Panel A) and European Commission - Directorate-General for Employment, Social Affairs and Inclusion (DG EMPL), labour market policy interventions database (Panels B and C).

The specific number of programme participants in specific programmes over time warrants some discussion. First, there was a very large increase in the number of individuals participating in the “employment incentives” programme in 2020. This increase is due to Lithuanian government’s introduction of measures to address the COVID-19 crisis. Second, many programmes experience considerable fluctuations from year to year. This is partly attributable to differences in funding across years, with participation in many programmes dependent on the availability of EU funds.

3.2. Future impact evaluations could augment the databases used

Despite the richness of the data on which the evaluation in OECD (2022[1]) draws, additions to the data used in this analysis could enrich future analyses which aim to examine similar research questions. Incorporating new sources of data would allow for expanding the scope of the research questions to examine additional, related research questions. These possibilities are discussed in turn in this section.

In terms of the key research questions examined in the OECD (2022[1]) evaluation– the effects of the ALMPs studied on occupational mobility, employment and earnings – two limitations to the data are notable. *First*, the employment data do not contain information on hours worked. This is worth bearing in mind when interpreting the results, particularly in terms of the outcomes relating to days worked and daily wages. For example, if participating in an ALMP increases the probability that an individual will become employed on a part-time instead of a full-time basis, this would produce a bias on the estimated results: actual hours worked would be lower than suggested by the observed days worked, whereas hourly wages would be higher than suggested by the observed daily wages. In practice, this may not be a large problem given the low prevalence of part-time work in Lithuania: in 2020, Lithuania’s part-time employment rate of

5.5% was one of the lowest in the OECD, where it averaged 16.7% (OECD, 2022[14])[4] *Second*, data on training do not include information on the target occupations of the training or a broader classification of the type of training offered. For this reason, examining whether individuals enter the occupations for which they underwent training is outside the scope of the analysis, as is examining the effects by type of training undertaken.

In the current project between the OECD and the EC, the data permitted a comparison of the earnings net of the direct LES expenditures associated with programme participation – the so-called subsidy amount (*subsidija* in Lithuanian). For vocational training, the subsidy amount includes the cost of training services (paid to the training provider) and a training scholarship (paid to the individual), as well as additional incidental costs associated with the training such as compulsory medical examinations. For employment subsidies, the subsidy amount corresponds to the amount paid to the employer (which covers only part of the employer's total costs associated with employing the individual). These results are useful for obtaining a general sense of whether the increase in earnings could conceivably offset the direct costs of a programme, but they are insufficient for an in-depth understanding of the wider costs and benefits of a particular programme.

Future impact evaluations could also conduct detailed cost-benefit analyses, which consultations with stakeholders have indicated would be of primary interest to policymakers. In order to facilitate detailed cost-benefit analyses, it would be useful to obtain more detailed information on:

- Actual means-tested social benefits received, including (from 1 June 2020) those paid for individuals after they begin working, as well as other means-tested benefits such as Compensations for heating costs, drinking water costs and hot water costs (*Būsto šildymo, geriamojo ir karšto vandens išlaidų kompensacijos*), Reimbursement of part of housing rental (*Būsto nuomos mokesčio dalies kompensacija*), and additional child benefit for low-income families raising one or two children (OECD, 2021[15]). This could also be supplemented by information on disability benefits.
- Household data and income information to establish eligibility for benefits. This would facilitate a counterfactual analysis to establish the potential level of means-tested benefits individuals would have received had they remained unemployed, as well as the level of into-work social benefits they would have been eligible to receive.[5]

Cost-benefit analyses could also include estimates of other social benefits. These include quantifiable benefits that cannot be expressed in monetary terms, as well as qualitative benefits that may be difficult to quantify (HM Treasury, 2022[16]). Such benefits include, for example, non-monetary benefits arising from individuals participating in the labour market, such as strengthening social ties and gaining a sense of independence or self-worth from doing meaningful work. Nevertheless, taking into account such costs and benefits can facilitate evidence-informed policymaking.

---

[4] Assuming that rates of part-time unemployment among jobseekers becoming employed are identical to the aggregate statistics would mean that the effect on hours worked would be correspondingly lower: the estimated effects of participation in ALMPs would be lower to reflect the fact that 5.5% of all individuals becoming employed were employed part time. However, the actual adjustment may need to be even larger: to the extent that part-time jobs are a stepping stone to full-time employment (Spector, 2022[40]), unemployed workers may be disproportionally more likely to move into part-time work than already employed workers, leading to even more of a bias in the results.

[5] Into-work benefits, paid to social benefits recipients who have been unemployed for at least six months, are paid for up to one year, with the amounts reduced after three months of employment (Ministry of Social Security and Labour, 2022[38])

Given Lithuania's high emigration rates, a final additional administrative data source that could be useful to include in future impact evaluations would be the population register. [6] Lithuania has a higher emigration rate than almost any OECD country, and a majority of those emigrating are neither employed nor registered as unemployed (OECD, 2018[17]). Such emigration data could be useful in impact evaluations in several ways. *First*, they could provide a more accurate assessment of outcomes in an impact evaluation such as the one in OECD (2022[14]). The OECD impact evaluation assumes jobseekers remained in Lithuania and that any effects on employment would be evident by employment in the SODRA employment data. However, accounting for jobseekers are moving abroad instead of becoming employed in Lithuania could affect the estimated effects of ALMP participation. *Second*, using administrative data on migration could help address a different set of questions than that analysed in the OECD (2022[14]) impact evaluation: whether specific social and labour market policies affect an individual's decision to move abroad. In such an impact evaluation, the outcome variable would be whether or not an individual migrates abroad and the "treatment" would be a specific policy or programme to which specific individuals or populations sub-groups were subjected. This could help inform whether social and labour market policies could be improved to encourage individuals to remain in Lithuania instead of moving abroad.

[6] In Lithuania, data on emigration may be directly available from the population register: OECD (2018[17]) argues that emigrants have a strong incentive to report their departure as soon as they leave Lithuania given that they are otherwise obligated to pay health insurance contributions. Combining data from additional registries can help augment this information by identifying individuals whose emigration can be inferred if they are not observed in any of the other registries for a reasonable amount of time.

4 The data linking and preparation process was relatively straightforward

This chapter discusses the process of linking, cleaning, and shaping the data files provided to the OECD by the Lithuanian authorities to make them ready for evaluation purposes used in OECD (2022[1]). It also compares the statistics calculated from the individual-level data with comparable statistics from publicly-available sources.

4.1. Few difficulties were encountered during data cleaning, shaping and linking

Creating datasets usable for evaluation purposes required linking several datasets at various different levels. The information from the various LES and SODRA databases was merged using pseudonymised individual identifiers; information about employers was linked with the pseudonymised firm-level identifiers. Linking the data files was relatively straightforward, although the data did need to be restructured to make them compatible and facilitate the analysis. The unit of observation differed across the data files, requiring some data manipulation to ensure that they could be compiled into a joint database for analysis. The following units of observation were present in the data provided: (1) the registration ID corresponding to an individual's unemployment spell, (2) the ALMP measure in which an individual participated, (3) employment spell data for individuals who had been unemployed at some point during the 2014-20 period, (4) monthly data relating to employment (for those individuals who had not been unemployed during the 2014-20 period), and (5) firm-level data recording changes in the firm registry.

In terms of the sequence of processing, each database first underwent a set of consistency checks and was then merged together with the most closely related database, undergoing another set of consistency checks before being merged with another database. The unemployment registry was merged first with the ALMP database, then with the employment database. After reshaping into monthly data, these were merged to the firm registry data.

4.1.1. Data cleaning work was relatively simple

Considering that cleaning the data presents a challenge for virtually all types of data analysis, the data used in the Lithuanian evaluation were relatively straightforward to process, with relatively few problematic records. Nevertheless, there were several difficulties faced in the evaluation in OECD (2022[1]) that are worth highlighting:

- *First*, missing data values for the occupation variable required some additional data work to copy the data from other records for the same individual. In instances of uninterrupted employment spells at the same employer (in many cases there were no gaps, but gaps of up to 16 days were construed as being uninterrupted), occupation codes were copied from other employment records with non-missing occupational codes. For employment spells preceding the onset of unemployment, data on previous occupation recorded in the unemployment registry were used to assign individuals occupations in case the value of this attribute was missing in the employment data.

- *Second*, using administrative data did not eliminate a small number of clearly unreasonable records, such as duplicated or overlapping records. This required conducting consistency checks to ensure that such records were either excluded (in the case of duplicates) or the relevant dates trimmed (in the case of overlapping records). In the case of overlapping records, the working assumption made was that spell begin dates were more accurate than spell end dates.

Given the relatively small number of problematic records and comparable data sources, comparing the official statistics on registered unemployment and ALMP participation yielded a very good correspondence (see Section 4.2). This is reassuring: it indicates that there are no unexplained sources of differences in the composition of the registered unemployed, and that the results of the OECD (2022[1]) impact evaluation can be considered representative of the entire population of registered unemployed. However, the results may not be generalizable if ALMPs were targeted to a wider set of individuals (including e.g. broad categories of individuals not registered with the PES in the period studied).

4.1.2. Data linking was done through pseudonymised identifiers

Data were linked using pseudonymised individual and firm identifiers which masked the identity of the underlying individuals and firms. The pseudonymisation was conducted by the Lithuanian authorities. The correspondence across the databases was generally excellent and linking them proved relatively straightforward. For example, of the 194 761 records for entry into ALMPs in the data provided, all except 23 records could be merged with the unemployment registry data. In the employment data for individuals who had been unemployed at some point during the 2014-20 period, only 0.15% of records contained missing values for the employer identifier, meaning they could not be linked to a specific employer.

The only minor challenge relating to data linking pertained to the firm-level identifiers in a sub-set of the employment data used to apply firm-level attributes. The sub-set of the data where this was an issue were those individuals who were never registered unemployed during the 2014-20 period (termed the "never-unemployed" below). These data were used for the descriptive analysis of the take-up of vocational training and employment subsidies as well as for the analysis of the displacement effects. Specifically, firm-level identifiers were not available for never-unemployed individuals who were self-employed or a small number of individuals who were employed by such an entity. This means that self-employed individuals are excluded, for example from the calculations on the attributes of firms participating in training through tripartite agreements. This issue was present in 3.2 million of the 45.5 million observations in the analysis dataset relating to the never-unemployed (which was generally at the level of one individual per month), thus affecting roughly 7% of this database.

Linking the databases on ALMP participation and employment outcomes with the registry of enterprises is useful for understanding the characteristics of firms who are employing individuals participating in ALMPs. For example, firms employing ALMP participants differ markedly by the sector of economic activity (Figure A A.1). For employment subsidies, the accommodation and food service sector stands out in terms of both total employment within the sector and the share of firms using this ALMP. Over ten percent of firms in this sector employed at least someone via employment subsidies at some point in the 2018-20 period. On average during this period, individuals on employment subsidies accounted for roughly 1.6 percent of employment. As a share of employment, agriculture and construction also stand out in their use of employment subsidies. For the vocational training with a tripartite agreement, the transportation sector stands out in terms of its high share of employment associated with this measure: roughly 0.2 percent of workers employed during 2018-20 were employed after engaging in vocational training with a tripartite agreement. This reflect a large share of tripartite agreements for individuals training to obtain commercial drivers licenses.

4.1.3. Data shaping required generating a panel data format

The most fundamental question faced in terms of shaping the data was how to combine data in spells format with data in panel format. Spells data record only dates in status, such as the date on which an individual registered as unemployed, or the start and end date of an ALMP measure: the LES data were recorded in spells format. Panel data record individuals' status in every period, for example every month: the employment data were (mostly) recorded in panel format.

The data used in OECD (2022[1]) recast all of the data in monthly panel format to ensure that information from all data sources could be combined. In part, this was a question of precision: it is more straightforward to convert data in which the timing is recorded more precisely (for example, the precise date) to a data format in which timing is recorded less precisely (for example, at the monthly level). However, having the data stored in a monthly panel is also more suitable for the econometric approaches described in Section 5.2.2. In practice, the econometric approaches often required a special wide panel dataset to be created, where – rather than the time variable being calendar month per se – the time variable effectively captured both time spent in unemployment (labelled m in Section 5.2.2) and the time elapsed since the start of the ALMP measure or the start of eligibility (labelled t).[7]

In addition, certain attributes in the unemployment registry required additional manipulations for them to be suitable for analysis. Data from the unemployment registry containing additional dimensions, which varied depending on the jobseeker, would ideally have been supplied in separate tables and linked to unemployment spells via the registration ID. A prominent example is the information on barriers to employment: many individuals had no barriers listed, but some had many (e.g. "Working persons up to 29 years", "Raising a child under the age of eight"). Given that the information was contained in a flat file, with one record per unemployment spell, this information had to be parsed from the data in order to be suitable for analysis.

Even after deciding to place the data in panel format, there remains the additional question around how to organise data such that the unit of observation is an individual unemployed person. This question arises because each individual may participate in more than one ALMP measure, have more than one concurrent employment, and have more than one spell in unemployment (and hence more than one LES registration ID). The option selected in the case of Lithuania was to retain information from all sources with the exception of concurrent employment spells (or those beginning and ending in the same month). In the case of concurrent employment spells, total earnings were aggregated at the monthly level and this value was used for the analysis pertaining to earnings. For the other attributes such as occupation, the "main job" was identified as that one with the largest share of earnings for that individual in that month.

4.2. Statistics calculated from the microdata closely match official statistics, suggesting evaluation results are representative of registered unemployed

One important potential advantage of using administrative data in impact evaluations is the possibility of having comprehensive coverage. Calculating estimates of population-level effects in evaluations using administrative data is rather straightforward compared to evaluations based on survey data, which require making assumptions about the comparability of individuals within specific groups to derive population-level

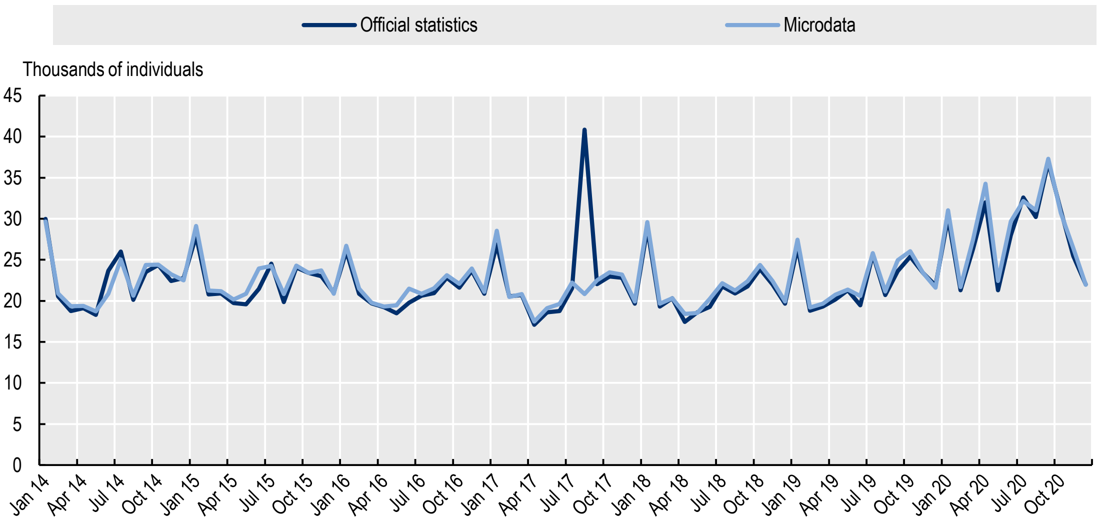
[7] This a special wide panel dataset was created from the panel dataset, with additional variables in the wide dataset relating to outcomes at different time periods relative to the reference point in the original panel dataset.

effects based on weights. Nevertheless, an important precondition for making simple population-level inferences using administrative data is to cross-check that the databases compiled are indeed complete.

In the case of the data used in this evaluation for Lithuania, statistics pertaining to registered unemployed calculated from the microdata closely match official statistics, suggesting evaluation results are broadly representative. Comparing monthly inflows into registered unemployment calculated from the microdata from those published by LES shows an almost perfect correspondence (Figure 4.1). Summed over the 2014-20 period, the discrepancy in total inflows between the two sources is 1.0%.

Figure 4.1. Unemployment registry inflow statistics from official statistics and microdata match almost perfectly

Monthly inflows into unemployment registry



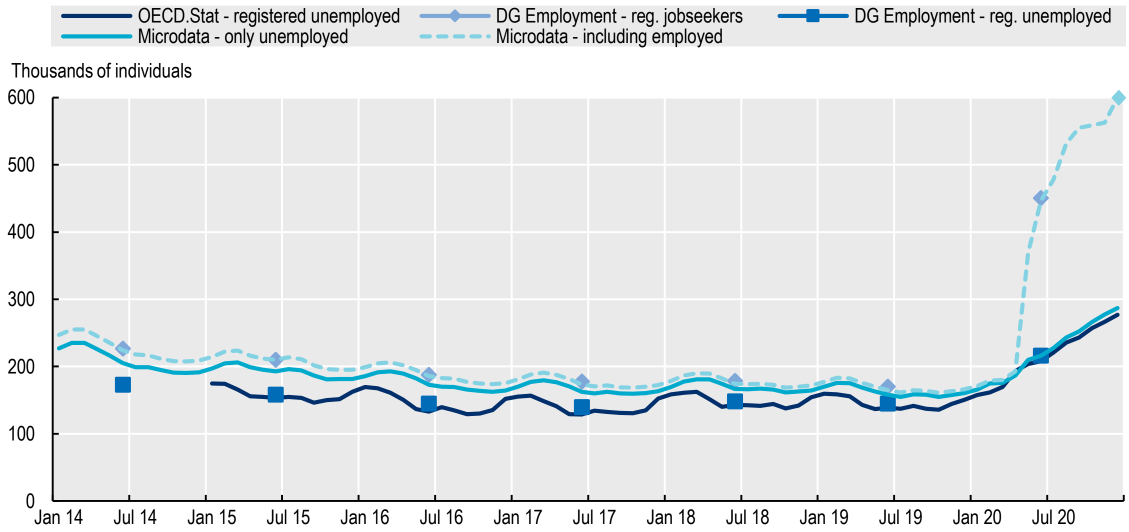
Note: Inflows from microdata are calculated by counting number of individuals with a registration date in a given calendar month and exclude those with registry status is denoted as "employed" (registry status code equal to 10).
Source: LES (2021[18]) and OECD calculations based on data from the Lithuanian Employment Service.

Comparing stocks of registered unemployment also shows a close correspondence between statistics calculated from the individual-level microdata and official statistics, although some differences are apparent due to differences in definitions. Figure 4.2 compares the monthly stocks of individuals registered with the LES from different sources and using slightly different definitions. Statistics from OECD (2021[19]) refer to individuals in the unemployment registry who are not also employed, while statistics in the EC database refer to all individuals.[8] The large discrepancy in the latter part of 2020 is attributable to schemes enacted in response to the COVID-19 crisis, when a large number of employed registered with the LES.

[8] The statistics from OECD (2021[19]) refer to the number of persons registered with the Employment Service who are neither in dependent employment nor self-employed (as defined in Paragraphs 1-2, Article 22, of the Law on Employment of the Republic of Lithuania).

Figure 4.2. Statistics on unemployment stocks calculated from the microdata closely match statistics from other sources

Monthly stocks of registered unemployed or jobseekers



Note: Stocks from microdata are calculated based on individuals registered at the beginning of the first day of the month. Statistics from European Commission (2022[20]) are annual averages of monthly statistics.
Sources: OECD (2021[19])), European Commission (2022[20]), and OECD calculations based on data from the Lithuanian Employment Service.

5 The methodology accounted for counterfactual outcomes

This chapter describes the research challenges faced in identifying the effects of the two ALMPs examined in OECD (2022[1]), the outcomes examined and the econometric approach used. It begins by outlining the scope of the evaluation and providing details on the outcomes measured, including the construction of the index used to measure occupational mobility. It then describes the challenges that need be addressed in order to accurately identify the effects and lays out the econometric approach. This includes considerations relating to minimising bias in the estimates such as accounting for unobserved heterogeneity between individuals in the treatment and comparison groups.

5.1. Deciding on what is being evaluated requires distinguishing between programme inputs, outputs and outcomes

Given that several types of evaluations of a programme can be conducted, it is useful to outline the scope of this counterfactual impact evaluation. For these purposes, it is useful to again apply the "results chain" framework (discussed also in the context of monitoring, in Section 2.2). A programme's results chain "sets out the sequence of inputs, activities, and outputs that are expected to improve outcomes and final outcomes" (Gertler et al., 2016[21]). In the context of the impact evaluation in OECD (2022[1]), a clear understanding of such a sequence is useful for clarifying the scope of the evaluation, as the ultimate focus is on evaluating outcomes and *not* on inputs, activities and outputs. The basic elements of the results chain, applied to the current evaluation, are as follows (Gertler et al., 2016[21]):

- Inputs: the resources available to the vocational training and wage subsidy programmes. These include LES staff and resources to administer the programmes, funds used for the vocational training or wage subsidies, and resources used by training providers (in the case of vocational training).
- Activities: activities that convert inputs into outputs of the project. These include the training programmes (in the case of vocational training) and the employment at the employers receiving the subsidies (in the case of employment subsidies).
- Outputs: tangible goods and services generated by programme activities. These include the number of individuals who completed training (for vocational training) or remained employed for at least a pre-specified duration (for employment subsidies) and certificates of training (in the case of some types of vocational training). Outputs are monitored by the LES and are at least indirectly under their control.
- Net outcomes (impacts): the effects that the programme achieves after the target population has received or been exposed to the programmes' outputs and activities, after taking into account counterfactual outcomes of participants had they not participated. These effects often reflect a change in the behaviour of the target population.

Counterfactual impact evaluations (CIEs) focus only on the last element of the results chain – the net outcomes. The elements proceeding it are discussed only to the extent that they help understand the outcomes. This makes it distinct from a process evaluation, which would examine whether programme activities have been implemented as intended. Furthermore, the outcomes are measured as *net* outcomes - after taking into account the counterfactual – and not *gross* outcomes as the current outcome indicators in the current ALMP monitoring framework in Lithuania (see Section 2.2). The gross outcomes would measure e.g. employment rates of participants without taking into account that some participants would have become employed anyway (for more details on the discussion of the counterfactual approach, see Section 5.2). The outcomes studied in the CIE conducted by the OECD are described in greater detail in the next sections.

5.1.1. Definitions of outcomes

In the case of Lithuania, the rich and comprehensive data available enable the analysis to track a wide set of outcomes in evaluating the programmes studied and over a relatively long period. The outcomes are tracked continuously over up to the three-year period starting with the beginning of the participation in a programme. Outcome values are calculated on a monthly basis and tracked over time relative to a reference month, which is defined either as the month when an individual enters an ALMP (for the treatment group) or that same calendar month for an individual in the comparison group who is matched to someone in the treatment group.

The research questions examined in the OECD (2022[1]) impact evaluation relate to labour market outcomes at time horizons from 3 to 36 months after entering the programmes. The following outcomes are examined:

- Probability of entering employment. This probability is measured using a binary outcome variable which is equal to 1 if individual is employed at certain time, and equal to 0 otherwise. The definition of employment includes various types of employment: by far the largest share of individuals are on regular, open-ended contracts, but it also includes becoming self-employed, individuals on fixed-term contracts, and other types of contracts (e.g. farmers).
- Cumulative employment duration. This measures the cumulative duration of all jobs held during the observation time, after the reference month. This measure is calculated on a monthly basis as the number of calendar days an individual was registered as employed based on observed employment spells. It takes without modifications instances where individuals were registered as employed for several days at a time, often in short-succession at a given employer.
- Occupational mobility. The analysis maps the occupation of individuals entering employment onto an occupational index, which can be interpreted as a “job ladder”. Nominal values of the index can be interpreted as based on constant 2015 prices. The construction of the index is detailed in Section 5.1.2.
- Cumulative earnings. This measures total earnings, gross of income taxes and employee contributions, in constant 2015 prices, received in all registered employment held during the observation time. Employer’s contributions are not included in this amount.
- Wages. This is calculated as earnings per calendar day, i.e. monthly earnings divided by the number of calendar days in the month during which an individual was employed in constant 2015 prices.[9] The number of calendar days an individual was registered as employed is calculated in the same manner as the cumulative employment duration. Due to data limitations, it is not possible to calculate an hourly rate.

[9] In months that individuals were not employed, their values are equal to missing. The results thus apply only to the subset of individuals employed at each specific point in time examined.

- Cumulative earnings net of subsidies. In order to assess the cost effectiveness of the studied ALMPs, the analysis also compares the benefits of each programme as expressed by cumulative earnings premium over the three year time horizon, gross of taxes and contributions, with the subsidy amount (*subsidija* in Lithuanian). The subsidy amount is a measure of the direct costs associated with the programme. For vocational training, the subsidy amount includes the cost of training services (paid to training provider) and a training scholarship (paid to the individual), as well as additional incidental costs associated with the training.[10] For employment subsidies, the subsidy amount corresponds to the amount paid to the employer for employing the individual. All amounts are in constant 2015 prices.

The following adjustments to the nominal values were made for the outcomes referring to monetary values (this includes wages, the occupational index, cumulative earnings, and cumulative earnings net of subsidies). Values for the period before the adoption of the euro in January 2015 are converted into euros according to the official conversion rate, at 3.4528 litas to the euro. The conversion into 2015 prices is done based on the monthly HICP index reported by Eurostat (2022[22]).

5.1.2. The occupational index was constructed based on observed wages

In addition to analysing outcomes typically examined in CIEs of ALMPs, such as employment probability or earnings, OECD (2022[1]) aims to address another important question: the effect of participation in ALMPs on occupational mobility. In order to provide a tractable measure of occupational mobility, the analysis relies on an occupational index, which is calculated from observed wages. Following the approach adopted by Laporšek et al. (2021[23]), a wage index is calculated for each detailed occupational code using data on the wages and employment of all employed individuals in Lithuania during the 2018-20 period. This index maps each of the 440 distinct occupational codes observed in the data into an index that has an intuitive and practical interpretation: an occupation whose index value is one unit greater than another occupation’s index value has an average real monthly wage that is one euro greater. Furthermore, increases and decreases in the index can be interpreted, respectively, as positive and negative changes in an individual’s occupation: climbing up or down the occupational ladder.

The analysis uses 4-digit ISCO-08 codes and is calculated from real monthly wages at constant 2015 prices. Given the different data format and content of the employment data for individuals depending on whether or not they were unemployed at some point during 2014-20, two separate indices are first calculated for each of the two groups, with employment-weighted averages of the two constructed to compile a single index. The precise calculation of the index is as follows:

1. Generate a measure of real prorated monthly income, using the data on earnings and calendar days employed, taking HICP inflation for Lithuania and using a base year of 2015. (Eurostat, 2022[22]).
2. Keeping only individuals on regular, open-ended employment contracts and excluding individuals with earnings below the statutory minimum wage (applied on a monthly level) and excluding outliers with extremely high wages (excluding the top 0.01% of the distribution).

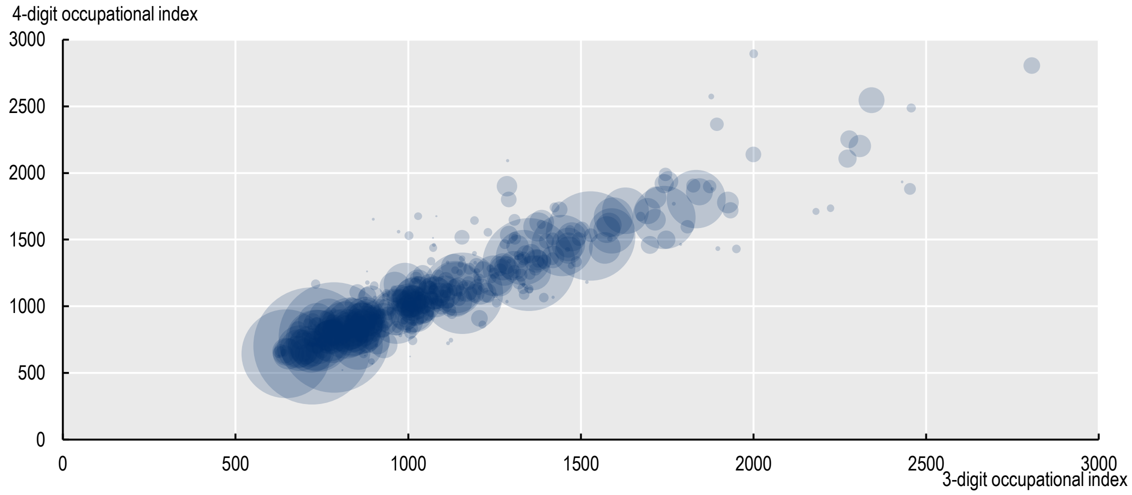
The procedure was repeated using occupation based on both 3-digit and 4-digit ISCO codes. While the analysis in OECD (2022[1]) impact evaluation uses 4-digit ISCO codes, the high correlation between indices calculated based on the two codes (with a correlation coefficient of 0.96) indicates that the choice of which index was is applied should not materially affect the results. Figure 5.1 plots the relationship between the two. The 4-digit code exhibits a greater number of index values at the upper end of the distribution, reflecting the fact that specialist occupations with high wages are not aggregated with lower-paying ones

[10] This may include cost of compulsory medical examinations and vaccination against communicable diseases (if applicable) as well as reimbursements of travel and accommodation costs.

to the same degree. On the other hand, the occupational index is only weakly correlated to the 1-digit ISCO codes, indicating that ISCO codes themselves would be insufficient proxies for occupational mobility (Figure A A.2).

Figure 5.1. Occupational indices based on 4-digit or 3-digit codes are similar

Occupational indices calculated at different levels of disaggregation

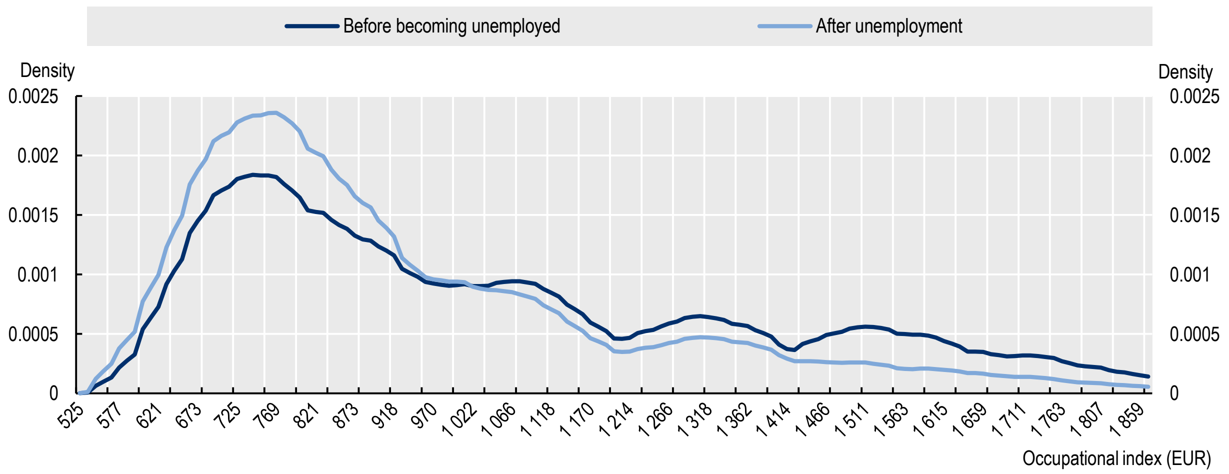


Note: Each shaded circle represents one 4-digit ISCO-08 code. The size of shaded circles is proportional to the number of individuals observed to be employed in the 4-digit occupation during the 2018-20 period.
Source: OECD calculations based on the Lithuanian Employment Service and Lithuanian State Social Insurance Fund Board.

The occupational index distribution for Lithuania shows that following unemployment, individuals who become re-employed disproportionally enter lower-ranked occupations (Figure 5.2). Following an unemployment spell, a larger share of individuals become employed in occupations whose mean monthly wages are below EUR 1 000; conversely, prior to becoming unemployed, a proportionally larger share of individuals were employed in higher ranked occupations. On average, individuals becoming re-employed have an occupational index that is approximately EUR 100 lower, corresponding with a roughly five percentage point drop down the distribution of the occupational index.

Figure 5.2. Individuals who become re-employed after unemployment disproportionally enter lower-paid occupations in Lithuania

Occupational index distribution before and after unemployment in Lithuania



Note: The heights of the lines indicate the relative share of individuals in occupations whose average wages are on the horizontal axis. The distributions are calculated for all individuals who were registered as unemployed at some point during the 2014-20 period. Observations with index value above EUR 1 866 are excluded from the kernel density chart.
Source: OECD calculations based on data from the Lithuanian Employment Service and Lithuanian State Social Insurance Fund Board.

5.2. The adopted econometric approach addressed several challenges in identifying the programmes' effects

This section first outlines the specific challenges that need to be addressed by impact evaluations of ALMPs in general and the Lithuanian programmes selected for evaluation in particular. It then discusses the details of the econometric approach used to identify the causal effects in the OECD (2022[1]) impact evaluation.

5.2.1. Challenges that needed to be addressed in order to accurately identify the effects

CIEs seek to measure the changes in outcomes experienced by programme participants, which can be attributed to the programme itself, as opposed to changes in external factors. For example, it may be observed that participants in an ALMP measure experience an improvement in their chances of being employed, but it is important to know whether the ALMP measure led to this improvement, rather than macroeconomic developments or changes in other policies.

CIEs aim to estimate programmes' effects (or "treatment effects") by comparing the actual outcomes of programme participants with what would have happened had those participants not in fact taken part in the programme. This "counterfactual" can never be observed in practice, so statistical or econometric techniques are needed to try and construct it. Normally, this is done by comparing the outcomes of participants (the "treatment" group) with a similar group of non-participants (the "comparison" group, particularly in the experiments also called the "control" group).

The "gold standard" in CIEs is often considered to be a randomised controlled trial (RCT), in which participation and non-participation in the programme is allocated randomly and the outcomes of these two (or more) groups are measured. Randomising participation in the programme minimises the chances that there are systematic differences between participants and non-participants, which are *not* related to participation in the project. For example, if participants are somehow more motivated or more capable than non-participants – which may arise if selection into a programme is voluntary rather than random – simply comparing participants and non-participants would produce biased estimates of the programme's treatment effects. RCTs should eliminate such sources of bias.

Despite these clear advantages, RCTs are not always suitable for evaluating ALMPs. Practically, it may be difficult to randomise participation (or even the timing of participation) in a programme: indeed, policymakers often seek to carefully target programmes to those most in need of support (for example, to those most at risk of long-term unemployment). Furthermore, randomising participation requires planning the evaluation in advance of launching a programme or intervention, may require monitoring to ensure compliance into the assigned treatment or control group, and may require keeping additional data to subsequently account for possible non-compliance in the analysis stage of the evaluation.

Given these factors, alternative techniques – which also compare participants and non-participants to estimate programme impact – are another tool for evaluating ALMPs. Unlike in RCTs, participation in the policy being evaluated is not randomly allocated in quasi-experiments. Instead, other methods are used to ensure that comparisons between participants and non-participants yield reliable estimates of programmes' treatment effects. For example, it may be possible to observe – and hence control for – any differences between participants and non-participants that would bias estimates of programme impact: this is the central tenet of matching methods and multivariate regression. Similarly, it may be possible to assume that comparing the changes experienced by participants and non-participants provides a reliable estimate of programme impact, even if it is not tenable to compare the levels of key outcomes variables. This difference-in-differences approach essentially compare the "value added" by a programme among participants and non-participants. Finally, it may be possible to exploit programmes' specific eligibility criteria. If individuals are only eligible for a programme above or below a precise cut-off in terms of some key variable (for example, age or household income), then participants and non-participants just either side of such a cut-off should be similar in all respects except their participation in the programme. Comparing such individuals – as is the objective of a regression discontinuity design – should therefore provide tenable estimates of programme impact (more discussion on applying the different evaluation methods to conduct CIEs of ALMPs can be found in OECD (2020[10])).

In the vocational training and employment subsidy programmes evaluated in OECD (2022[1]), participation was not randomly assigned and there were no strict eligibility criteria. The Lithuanian Law on Employment specifies the target groups of unemployed individuals that are to be the primary beneficiaries of ALMPs (Republic of Lithuania, 2016[24]). Guidelines specify which programmes should be preferentially applied to each of the specific groups (Ministry of Social Security and Labour, 2017[25]). These specify, for example, that workers with disabilities are to receive a wide array of support via training and subsidies relating to employment, including ones dedicated to supported employment and rehabilitation (which are not the subject of the OECD (2022[1]) impact evaluation). For this reason, individuals with disabilities who have an assessed working capacity 25% or less are not included in the subsequent analysis.

Taking into account the above guidelines, it is important to note that the law gives LES counsellors the right to exercise discretion in deciding whether to refer an individual to a specific measure. LES counsellors appear to exercise this discretion in practice. To the extent that these target groups can be accurately gleaned from the administrative data, a sizable share of individuals do not meet any of the above criteria in practice. Among the variables coded in the data by LES counsellors, the highest shares of participants fulfil the age criteria for target groups (Table 5.1).

Table 5.1. Age criteria are most prominent for participation in vocational training and employment subsidy programmes

Share of individuals participating in either the vocational training or employment subsidy programmes belonging to specific target groups

Target group	Share of vocational training participants fulfilling criterion	Share of employment subsidy participants fulfilling criterion
Persons up to 29 years of age	39.9	37.6
Persons over 50 years of age	23.3	39.5
Raising a child under the age of eight	11.8	10.0
Unskilled unemployed (individuals without professional qualifications or education recognised in Lithuania)	5.4	5.8
Working-age persons with a disability and a level of working capacity of 45–55 percent	2.8	4.3
Unemployed persons starting work for the first time after having acquired their current qualifications	2.8	1.9
Working-age persons with a disability and a level of working capacity of 30–40 percent	0.7	2.3
Other categories	0.0	1.5

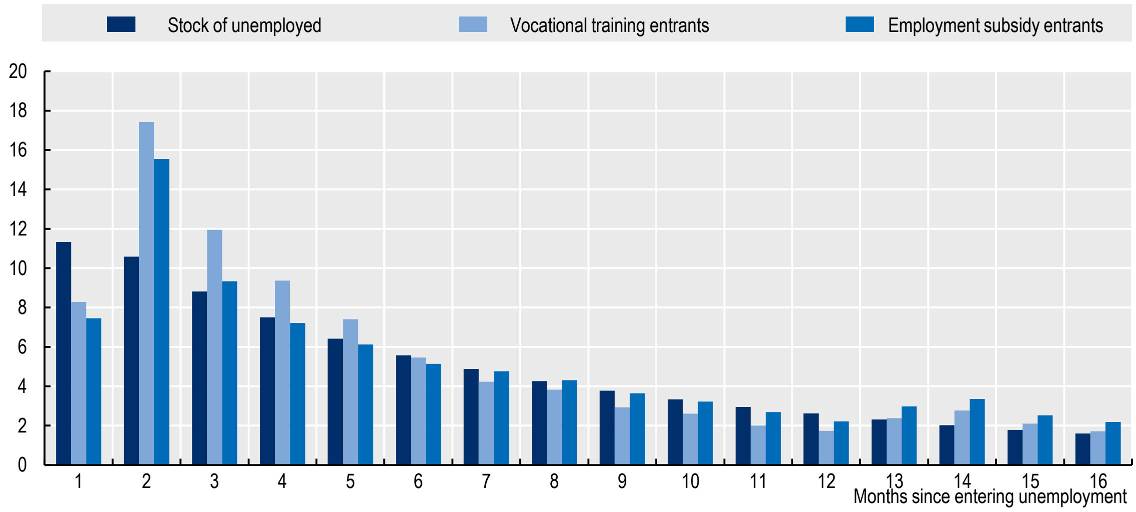
Note: Categories are not mutually exclusive. Values are as coded in the individual-level database. Values are calculated based on the attributes of all individuals who participated in either the vocational training or employment subsidy programmes at some point during the 2014-20 period.
Source: OECD calculations based on data from the Lithuanian Employment Service.

In practice, different subgroups of individuals also entered different types of vocational training programmes, which provided a further impetus for doing sub-group analyses. Examining the most commonly entered vocational training programmes shows considerable differences between men and women (Annex Table 4), including after accounting for whether individuals entered the training via tripartite agreements. Among women, for example, the most commonly entered programmes were programmes for manicurists and accountants (among entrants with and without tripartite agreements, respectively). For men, the most commonly entered programmes were those for obtaining commercial motor vehicle licenses.

Another challenge in identifying the effects of participation relates to the importance of duration in unemployment in determining which individuals selected into the programmes. In practice, individuals enter vocational training or the employment subsidy programme at various points in their unemployment spell, but the exact timing is not predetermined and varies substantially (Figure 5.3). Individuals disproportionally enter vocational training or employment subsidies at the beginning of their unemployment spells and, to a lesser extent, after being unemployed for 13-17 months. In practice, individuals generally enter vocational training after being unemployed for 3-5 months for formal training and after 2-3 months for non-formal training, although individuals can enter training immediately after becoming unemployed. For employment subsidies, exact time of entry varies also considerably in practice (the median unemployment duration at entry is five months).

Figure 5.3. Individuals disproportionally enter vocational training or employment subsidies at the beginning of their unemployment spells

Share of all individuals among registered unemployed, individuals entering vocation training, or individuals beginning employment subsidies



Note: The distributions are calculated for all individuals who were registered as unemployed at some point during the 2014-20 period. For the "All unemployed" category, shares were calculated based on monthly unemployment stocks. Individuals unemployment durations longer than 18 months are excluded from the chart. Duration of unemployment is rounded up to the nearest month.
Source: OECD calculations based on data from the Lithuanian Employment Service.

Given these patterns in the timing of ALMP starts, making simple comparisons between those who did and did not participate in the ALMPs studied will not yield reliable estimates of their impact on labour market outcomes. This is because many of the non-participants do not participate simply because they are able to find a job quickly and independently (and hence exit employment) without necessarily receiving support from the LES. Such individuals may have better labour market outcomes than ALMP participants by construction: having exited unemployment quickly, they have an improved chance of keeping their new job, giving them a better chance of being employed and having higher earnings in the future. Such individuals may also be more motivated or able than ALMP participants in the first place, in ways that may be difficult to observe in the data. Additionally, in the case of vocational training, non-participants will not have their unemployment spells extended mechanically by being “locked-in” to training.

One potential problem often encountered in impact evaluations of ALMPs concerns the question of how to deal with multiple, sequential entries into ALMPs. In the presence of multiple interventions and possible overlap between different ALMPs, identifying the precise effects of one specific ALMP presents an important challenge. In the case of Lithuania, this is not a major concern: during the vast majority of unemployment spells (92.5%), individuals entered into only one ALMP during their entire unemployment spell. In the remaining 7.5% of cases, individuals entered either two or three ALMPs in total. Of the latter, a sizable proportion involved short ALMP durations (e.g. less than one month). For the purposes of the evaluation, we focus on the first ALMP entered during an unemployment spell. The exception to this rule concerns the cases where individuals enter an ALMP for less than a month and enter another ALMP within the same month. In this case, we examine the effect of the ALMP which is of longer duration.

5.2.2. The econometric approach combined dynamic coarsened exact and propensity score matching

To respond to these challenges, the evaluation of the programmes studied focusses on subsets of individuals that have endured a set number of months in unemployment and compares the labour market outcomes of those who begin the ALMP in that month with those who are still "waiting" for support from an ALMP measure or for some other way out of unemployment. In particular, this "dynamic selection-on-observables" approach – developed by Sianesi (2004[26]) – hinges on estimating (and then aggregating) separate treatment effects for each pre-treatment duration (m, the amount of time between registration and entry into the ALMP measure studied) and the time horizon of interest (t, the amount of time elapsed since the start of the ALMP measure, when labour market outcomes are measured). The *potential* labour market outcomes (such as employment or earnings) for an individual (i) can be written Y_{imt}^{d}, where $d = 1$ under treatment and $d = 0$ otherwise.[11] The average treatment effect on the treated ($D_{im} = 1$) for each t can then be written:

$$\gamma_t = E[Y_{imt}^{1}|D_{im} = 1] - E[Y_{imt}^{0}|D_{im} = 1]$$

In this framework:

- The treatment group comprises those individuals who *begin* treatment in period m.
- The potential comparison group comprises individuals who were still unemployed in period m, but who did not enter an ALMP (including those that went on to start training in future periods).

The dynamic selection-on-observables approach therefore recasts the simple comparison between participants and non-participants – which would lead to biased estimates of programmes' treatment effects – as a series of more reasonable comparisons between individuals with similar experiences of unemployment. The approach also tackles the issue of multiple treatments by focussing on the *first* ALMP measure that individuals receive and then effectively treating any subsequent measures as part of individuals' labour market outcomes. Thus, if an individual stays unemployed so that they can enter another ALMP (or enters another ALMP because they have remained unemployed), this is treated as information about outcomes rather than information about subsequent treatments.

A wide array of variables are used in the propensity score calculations. These variables can be grouped into the following categories:

- **Prior labour market histories:** occupational index of last occupation, cumulative earnings, cumulative days worked, as well as binary variables to indicate individuals without prior employment histories.
- **Demographic characteristics:** age, gender, and education.
- **Skills levels:** English language skills, basic IT skills and clerical skills.
- **Motivational factors:** assessed motivation to become employed (as determined by LES counsellors), level of unemployment benefits, which may affect the willingness of an individual to search for or accept a job as well as the willingness to engage in training.
- **Barriers to employment** (as categorised by the LES): age under 30 or over 50, raising a child under the age of 8, unskilled employed, and other barriers.

Some of the variables in the raw LES data required recoding before they could be used in the propensity score calculations. This was the case for variables which were originally coded as categorical variables such as counsellor-assessed jobseeker motivation, which was assessed on a three-point scale (low,

[11] To facilitate the description of the identification approach, the exposition in this section describes an approach which does not employ difference-in-differences to account for unobserved heterogeneity. A similar framework can be adapted to the approach difference-in-differences approach used in the OECD (2022[1]) impact evaluation.

medium, high). Such variables were recoded into separate binary variables (in the case of the jobseeker motivation variable, into three separate variables for each value, as well as one binary variable indicating whether this variable had a missing value). Summary statistics of the resulting variables are included in Annex Table 5 through Annex Table 8.

These extensive set of covariates, which include prior labour market histories, are arguably sufficient to control for important unobserved characteristics which likely affect selection into treatment. Numerous studies, including the seminal work of Heckman et al. (1998[27]), find that including detailed information on recent labour force histories and earnings can mitigate the effects of unobserved characteristics in constructing comparison groups. Caliendo, Mahlstedt and Mitnik (2017[28]) provide evidence that usually unobserved aspects such as personality traits, social networks and life satisfaction played an important role in determining selection into ALMP treatment in Germany, but that conditioning on detailed labour market histories can implicitly capture much of the information contained in the usually unobserved variables. Furthermore, the analysis for Lithuania was able to incorporate aspects typically not available from administrative registries, based on the motivation of the jobseeker as assessed by the employment counsellor. The validity of the non-experimental approach is further bolstered by the general finding in the meta-analysis of ALMPs in Card, Kluve and Weber (2018[29]), who find that results from experimental and quasi-experiment research designs do not systematically vary in terms of finding “less positive” or “less significant” results.

An additional step of the procedure of generating comparable treatment and control groups involves imposing the requirement that individuals are matched to individuals with identical characteristics for several key variables. This exact matching is applied, based on monthly data, using the following variables:

- **Age**: under 30, 30-50, over 50;
- **Unemployment benefit receipt**: individual is receiving unemployment benefits in a given calendar month, individual is *not* receiving unemployment benefits in a given calendar month; and
- **Unemployment duration**: 2 months or less, 2-3 months, 4-6 months, 7-11 months, 12-23 months, 24 months or more.[12]

Applying the rich set of variables used in the propensity score matching, combined with the exact matching on certain characteristics, results in treatment and comparison groups that arguably meet a fundamental requirement for identification, conditional independence – that treatment assignment is independent of potential outcomes. This is confirmed by placebo tests for periods prior to treatment, which yield statistically insignificant results for most of the observation periods examined (

[12] In practice, given the collinearity of the unemployment duration variable with the variables on age and unemployment benefit receipt, the unemployment duration variables were aggregated for some specific combinations. This accounted for the fact that there were small numbers of e.g. jobseekers over 50 not receiving unemployment benefits upon becoming unemployed.

Figure A A.3 and Figure A A.4). Nevertheless, given that the results are not insignificant for *all* the periods, a difference-in-difference estimator is employed for all the estimates presented in the OECD (2022[1]) impact evaluation. Tests of the balancing properties before and after matching are provided in Annex Table 5 through Annex Table 8. These indicate that the balancing properties are broadly satisfied.

5.2.3. Minimising bias in the estimates: Using a difference-in-differences approach to control for unobserved heterogeneity, a nearest-neighbour treatment effect estimator and obtaining correct standard errors

A fundamental assumption in propensity score matching is that all differences between programme participants and non-participants relevant for outcomes are captured in their observed characteristics. This means that after conditioning on observed covariates selection into the treatment can be considered random (Imbens, 2000[30]). In the case that selection into programme is governed not only by observable but also *un*observable individual characteristics that are correlated with the potential outcomes, then propensity score matching provides biased estimates of treatment effects. In order to address this issue, the analysis combines propensity score matching with a difference-in-differences approach (Heckman et al. (1998[27]), Smith and Todd (2005[31])). This approach compares intertemporal changes in outcomes between participants with changes in outcomes for the comparison group, with changes measured relative to a pre-programme benchmark period:

$$\Delta_{im}^{\tau}= \left(y_{im,t=\tau}^{1} - y_{im,t=-\tau}^{1}\right) - \left(y_{j,t=\tau}^{0} - y_{j,t=-\tau}^{0}\right)$$

where y is the actual outcome of interest, $y_{im,t=\tau}^{1}$ and $y_{im,t=\tau}^{0}$ are post-unemployment outcomes for participants for controls, respectively, and $y_{im,t=-\tau}^{1}$ and $y_{im,t=-\tau}^{0}$ are their pre-programme outcomes. By controlling for time-invariant unobserved heterogeneity, this specification ameliorates the bias resulting from unobserved differences between participants and non-participants.

Many different types of estimators could conceivably have been employed for estimating the average treatment effect. These include parametric estimators (such as doubly-robust estimators), inverse probability weighting estimators, matching estimators, and kernel matching estimators. These estimators each have various benefits and drawbacks, with the question of which estimator performs best in empirical applications having been examined in several studies without an unambiguous answer. For example, Frölich (2004[32]) finds that a version of kernel-matching based on local regressions with finite sample adjustments performs best in Monte Carlo simulations with up to 1,600 observations. By contrast, Busso et al. (2014[33]) find that in Monte Carlo simulations and empirical data-generating processes, inverse probability weighting using normalized weights exhibits the best properties in finite samples where overlap is good. They also find that in instances of data-generating processes with poor overlap, bias-corrected matching with a fixed number of neighbours is the most effective. Huber et al. (2013[34]) perform an empirical Monte Carlo study on a large number of estimators measuring average treatment effects on the treated and find that no estimator is superior in all designs and for all outcomes and that bias-adjusted radius matching estimators perform best on average. The relative performance of estimators arguably depends strongly on features of the data-generating process (Busso, DiNardo and McCrary, 2014[33]) which is unknown to the empirical researcher in practice.

Given the lack of a clear favourite for the estimator and the relatively large sample sizes for the programmes to be evaluated in Lithuania, the OECD (2022[1]) impact evaluation employs a nearest neighbour direct matching estimator. Huber et al. (2013[34]) note that such a matching estimator exhibits the smallest bias for all sample sizes: using only the most similar controls reduces bias and exhibits robustness to propensity score misspecification because it remains consistent if the empirical propensity score specification is a monotone transformation of the true model. One the other hand, it has the drawback of being inefficient (i.e., having higher variance than other estimators). Nevertheless, with large sample sizes the superior properties in terms of bias are relatively more important, with variances tending to zero asymptotically.

A final point concerns the calculation of the standard errors of the estimates. Given that propensity scores are estimated and not known in observational studies, estimating bootstrapped standard errors has been a commonly-used approach in propensity score matching. However, as shown by Abadie and Imbens (2008[35]) the standard bootstrap is not generally valid for matching estimators. In a later paper (Abadie and Imbens, 2016[36]), they derive an analytical solution to address the issue, proposing an adjustment to the large sample variances of propensity score matching estimators for estimating the average treatment effect on the treated. The standard errors in this analysis are estimated using the adjustment outlined in that paper. As Abadie and Imbens (2016[36]) show, the sign of the adjustment depends on the data generating process, and ignoring the estimation error in the propensity score may lead to confidence intervals that are either too large or too small. In practice, in the Lithuanian results, applying the adjustment results in narrower confidence intervals.

6 Suggested impact evaluation framework with roadmap

A rough outline of the key elements to consider when evaluating a specific policy or programme is provided below. This should not be seen as a linear process as many of these elements are interdependent or some of them may not be relevant in certain contexts. The below list should rather be regarded as a checklist of the main aspects, rather than a step-by-step roadmap to follow.

Phase 1: Understanding the needs for counterfactual impact evaluations

- **Choose whether/which programme or intervention to evaluate, involving all relevant stakeholders.** Provide a very initial rough assessment of the feasibility, the possible costs and the likely benefits of conducting an evaluation of the respective intervention.

  The analysis units in the LES and the Ministry of Social Security and Labour (*Socialinės apsaugos ir darbo ministerija* – SADM) to coordinate the process of choosing the interventions to be evaluated, involving all relevant stakeholders, most importantly policymakers.
- **Establish the evaluation purpose.** Assess whether the main purpose is to ensure accountability for stakeholders, comparing the relative merits of different programmes, learn about the programme results for future policy design (including targeting of measures to sub-groups), or something else.

  The analysis units in LES and SADM to coordinate discussions on the purpose of the evaluation, involving all relevant stakeholders, most importantly policymakers.
- **Decide the evaluation criteria.** Given the evaluation purpose, which are the main criteria to determine success of the intervention (e.g. according to OECD criteria, such as relevance, coherence or effectiveness).

  The analysis units in the LES and the SADM to coordinate the discussions on the evaluation criteria, involving all relevant stakeholders, most importantly policymakers.
- **Determine the evaluation approach.** Given the evaluation criteria, decide about the suitable evaluation approach (e.g. process evaluation, impact evaluations (CIE), cost-benefit evaluation).13

  The analysis units in the LES and the SADM to coordinate the discussions on the evaluation approach, involving all relevant stakeholders, most importantly policymakers.

Phase 2: Planning counterfactual impact evaluations

- **Determine the outcomes of interest.** This step involves the decision what will be measured to determine success regarding the key evaluation questions (results indicators). These indicators should describe the intervention's overall objectives and results in operational and measurable terms.

[13] This step may come to the conclusion that other evaluation approaches are more suitable in specific context. In line with the focus of this note, the following describes the steps for conducting a (counterfactual) impact evaluation.

The analysis units in the LES and the SADM to coordinate the discussions on the outcomes of interest, involving all relevant stakeholders, most importantly policymakers and experts who can opine on the feasibility of obtaining data on these outcomes.

- **Determine a Theory of Change.** The Theory of Change should provide a clear indication of the main underlying assumptions of the intervention design. The goal is to define questions that can be answered with the CIE. Evaluation questions typically depend on the chosen evaluation criteria and thus the purpose of the evaluation. For impact evaluations, the basic evaluation question is typically "What is the causal effect (or impact) of a programme on an outcome of interest?"

  The analysis units in the LES and the SADM to coordinate the discussions on the evaluation questions, involving all relevant stakeholders, most importantly policymakers.

- **Assess ethical considerations.** Ensure that ethical standards will be fulfilled in the design and implementation of the evaluation and related data collection.

  The analysis in the LES and the SADM to coordinate the discussions on the ethical considerations, involving all relevant stakeholders, most importantly policymakers and legal units.

- **Establish an evaluation team.** The evaluation team should include a range of well-formed individuals with expertise in both the specific intervention, policy and evaluation methods.

  The analysis units in the LES and the SADM to decide whether to conduct the evaluation in-house or contract out. In the first case, the choice of the evaluation team is the internal decision of the analysis units. In the latter case, the analysis units draft the procurement documents (defining what, when and how to evaluate) with the support of experts of public procurement in the LES and the SADM. In case a CIE is contracted out, the next steps are conducted in cooperation between the analysis teams and the contractor (the winner of the procurement process). The analysis teams will be supervising the process, while the contractor will implement the actual evaluation. (Note that for simplicity, the activities in subsequent steps are outlined under the assumption that the evaluation is conducted in-house).

- **Choose an impact evaluation method.** The key step in the design is to determine the (possibly multiple) evaluation methods that are feasible and provide credible results. The chosen method should be derived from the operational rules of the programme (e.g. determining whether an experimental study is feasible).

  The chosen evaluation teams in the LES and the SADM to discuss and decide the evaluation methods.

- **Determine the eligible population and sampling frame.** To (randomly) select treatment and comparison groups, one needs to delineate clearly the full eligible population from which groups are drawn. Once the eligible group is defined, a next step is to establish a "sampling frame", which is a comprehensive list of all individuals in the eligible population.

  The chosen evaluation teams in the LES and the SADM to discuss and determine the eligible population and sampling frame.

- **Plan the data collection**

  Decide on using survey and/or administrative data. If administrative data are available, these are often preferable since these data will typically involve lower effort and costs. However, if administrative data do not provide information on the chosen outcomes of interest, additional survey data may be required.

  Choose the sample size (including power calculations). If survey data are collected, it may not be feasible to collect data covering the entire treatment and comparison groups. In this case, power calculations are a key step to determine ex-ante whether the foreseen impact evaluation design will be able to provide valid results given the sample size.

The chosen evaluation teams in the LES and the SADM to plan the data collection, involving other units and stakeholders (for example other Ministries, the Tax Authority) as well as data protection delegates if necessary.

- **Assess potential risks to the evaluation.** Assess which factors may impede the internal validity of the evaluation design (e.g. selective survey attrition, spill-over effects, unintended behavioural effects…).

  The chosen evaluation teams in the LES and the SADM to assess potential risks to the evaluation.

Phase 3: Implementing and managing impact evaluations

- **In case sampling is needed (e.g. entire datasets are too big):** Draw the evaluation sample. In case the evaluation is not conducted with the entire treatment and comparison group, the evaluation sub-sample should be (randomly) drawn from the full eligible list.

  The chosen evaluation teams in the LES and the SADM to decide on the needs and methods to draw samples.

- **In case of RCTs: Conduct randomisation.** When an experimental evaluation is feasible, randomisation must be conducted prior to the intervention start. The randomisation may be administered in collaboration with programme managers and staff, but should be monitored by the evaluators. 14

  The chosen evaluation teams in the LES and the SADM to randomise the eligible group and communicate the randomisation set-up clearly to policy implementers, generating awareness and ownership of the evaluation. The evaluation teams conduct also training to policy implementers to apply randomisation properly if needed. The evaluation teams in in the LES and the SADM subsequently monitor that their proposed randomisation is followed by the implementers and intervene if necessary (discuss with the implementers the challenges and ways to overcome these). Applying randomisation could be supported by small-scale IT developments to the main operational IT systems, applying internal controls that ensure the application of randomisation. The input on the IT development should be provided by the evaluation teams in conjunction with the programme managers. The IT developments should be implemented by the IT units in the LES and the SADM.

- **Collect survey and/or administrative data.** This involves collecting either baseline and/or already follow-up data, depending on the chosen evaluation method and type of data.

  The LES and the SADM analysis units to draw the data needed for the evaluations from the internal registers of the ministries and make additional data requests to other stakeholders in case necessary (e.g. external national registers), involving IT units and data protection delegates in the process in case needed.

- **Clean and process data.** This step is needed in most cases both for administrative and survey data. Typical steps involve cleaning the data from erroneous observations, combining various sources of data, and checking for the validity of reported information.

  The chosen evaluation teams to process the data.

- **Analyse data.** The final step often involves three key parts, although more steps might be needed dependent on the analysis at hand: 1) test whether the main assumption of the chosen research design are fulfilled (e.g. baseline balance); 2) run the empirical analysis to estimate treatment

[14] Optionally, to increase the credibility of the impact evaluation results and increase the visibility of the subsequent results (assuming they are published), at this stage the RCT may be registered in a trial registry which outlines key features of the trial and the proposed evaluation of the outcomes (a pre-analysis plan). The latter can increase the credibility of the findings and address questions of specification searching by setting out the expected analysis approach in advance (Banerjee et al., 2020[37]). One example of an RCT registry is the one maintained by the American Economic Association: https://www.socialscienceregistry.org.

effects; and 3) check how sensitive the estimated effects are to deviations from the underlying assumptions.

The chosen evaluation teams to analyse the data.

- **Draft the report(s) on the evaluation results.** The content and writing style of the report should consider the target audience (e.g. the description of methodology can be more detailed and technical when a report is targeting researchers, but focussing on key messages when targeting high-level policymakers).

  The chosen evaluation teams to draft the reports, seek feedback from all relevant stakeholders (particularly policymakers, ideally also from external researchers) and incorporate relevant comments in the reports.

Phase 4: Disseminating results, ensuring policy uptake and managing knowledge

- **Disseminate internally.** This involves discussing the evaluation results with key policymakers, programme managers and staff, as well as supporting policymakers to translate the evaluation results into action plans, redesigned policies and guidelines for policy implementers.

  The chosen evaluation teams to present the reports to their colleagues (including policymakers and policy implementers) and support the policymakers to translate the results into action plans, redesigned policies and guidelines.

- **Disseminate externally.** The evaluation results should generally be made available to the civil society, irrespective of whether the initial evaluation purpose was for accountability or learning. This may include distilling the most important findings, publishing them and possibly involving designing an outreach strategy.

  The chosen evaluation teams to set up a dissemination strategy (preferable already during the planning phase of the evaluation) with the support of LES and SADM communication experts. After finalisation of the report and internal discussions, the report is disseminated according to the dissemination strategy by the LES and SADM evaluation teams in cooperation with the communication experts.

- **Learn for future evaluations.** This involves drawing lessons from the evaluation regarding internal analytical capacity, data availability and process.

  The chosen evaluation teams in LES and SADM to reflect on the evaluation process after its completion, analysing by each step, which approaches were successful and which steps were challenging. Based on these reflections, the evaluation teams in LES and SADM draw up an action plan to improve the evaluation process for the next evaluations (tasks for the analysis teams as well as proposals for other units, such as the IT units and data protection delegates).

References

A. Smith, J. and P. E. Todd (2005), “Does matching overcome LaLonde’s critique of nonexperimental estimators?”, *Journal of Econometrics*, Vol. 125/1-2, pp. 305-353, https://doi.org/10.1016/j.jeconom.2004.04.011. [31]

Abadie, A. and G. Imbens (2016), “Matching on the Estimated Propensity Score”, *Econometrica*, Vol. 84/2, pp. 781-807, https://doi.org/10.3982/ecta11293. [36]

Abadie, A. and G. Imbens (2008), “On the Failure of the Bootstrap for Matching Estimators”, *Econometrica*, Vol. 76/6, pp. 1537-1557. [35]

Banerjee, A. et al. (2020), *In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plans for RCTs in Economics*, https://economics.mit.edu/files/19395 (accessed on 25 May 2022). [37]

Busso, M., J. DiNardo and J. McCrary (2014), “New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators”, *Review of Economics and Statistics*, Vol. 96/5, pp. 885-897, https://doi.org/10.1162/rest_a_00431. [33]

Caliendo, M., R. Mahlstedt and O. Mitnik (2017), “Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies”, *Labour Economics*, Vol. 46, pp. 14-25, https://doi.org/10.1016/j.labeco.2017.02.001. [28]

Card, D., J. Kluve and A. Weber (2018), “What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations”, *Journal of the European Economic Association*, Vol. 16/3, pp. 894–931, https://doi.org/10.1093/jeea/jvx028. [29]

European Commission (2022), *Persons registered with Public Employment Services database*, Directorate-General for Employment, Social Affairs and Inclusion, https://webgate.ec.europa.eu/empl/redisstat/databrowser/product/page/LMP_RJRU$TPS00081 (accessed on 24 May 2022). [20]

European Commission (2020), *How to use administrative data for European Social Funds counterfactual impact evaluations : a step-by-step guide for managing authorities*, Publications Office, https://data.europa.eu/doi/10.2767/721497 (accessed on 11 October 2020). [12]

European Commission (2019), *Advanced counterfactual evaluation methods : guidance document.*, Publications office, https://doi.org/10.2767/464242. [11]

European Commission (2018), *Programming period 2014-2020. Monitoring and Evaluation of European Cohesion Policy. European Social Fund. Guidance document*, https://ec.europa.eu/sfc/en/system/files/ged/ESF%20monitoring%20and%20evaluation%20gui [3] dance.pdf.

European Commission, Directorate-General for Employment, Social Affairs and Inclusion (2020), *Counterfactual impact evaluation of European Social Fund interventions in practice : guidance document for managing authorities*, Publications Office, https://data.europa.eu/doi/10.2767/55495. [13]

Eurostat (2022), *HICP - monthly data (index)*, https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=prc_hicp_midx&lang=en. [22]

Frölich, M. (2004), "Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators", *Review of Economics and Statistics*, Vol. 86/1, pp. 77-90, https://doi.org/10.1162/003465304323023697. [32]

Gertler, P. et al. (2016), *Impact Evaluation in Practice, Second Edition*, Washington, DC: Inter-American Development Bank and World Bank, https://doi.org/10.1596/978-1-4648-0779-4. [21]

Heckman, J. et al. (1998), "Characterizing Selection Bias Using Experimental Data", *Econometrica*, Vol. 66/5, p. 1017, https://doi.org/10.2307/2999630. [27]

HM Treasury (2022), *The Green Book: Central Government Guidance on Appraisal and Evaluation 2022*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063330/Green_Book_2022.pdf (accessed on 17 June 2022). [16]

Huber, M., M. Lechner and C. Wunsch (2013), "The performance of estimators based on the propensity score", *Journal of Econometrics*, Vol. 175/1, pp. 1-21, https://doi.org/10.1016/j.jeconom.2012.11.006. [34]

Imbens, G. (2000), "The role of the propensity score in estimating dose-response functions", *Biometrika*, Vol. 87/3, pp. 706-710, https://doi.org/10.1093/biomet/87.3.706. [30]

Laporšek, S. et al. (2021), "Winners and losers after 25 years of transition: Decreasing wage inequality in Slovenia", *Economic Systems*, Vol. 45/2, p. 100856, https://doi.org/10.1016/j.ecosys.2021.100856. [23]

LES (2021), *Statistiniai rodikliai [Statistical indicators] - tab Laikotarpio rodikliai, indicator Registruota bedarbi*, https://uzt.lt/darbo-rinka/statistiniai-rodikliai/ (accessed on 27 April 2021). [18]

Lietuvos Respublikos socialinės apsaugos ir darbo ministerija (2017), *Dėl Darbo rinkos stebėsenos sąlygų ir tvarkos aprašo patvirtinimo*, https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/f4bb4fc0578611e78869ae36ddd5784f. [4]

Ministry of Social Security and Labour (2022), *Piniginė socialinė parama nepasiturintiems gyventojams (socialinė pašalpa, būsto šildymo, geriamojo ir karšto vandens išlaidų kompensacijos) [Cash social assistance for the poor (social benefits, compensation for heating, drinking and hot water)]*, https://socmin.lrv.lt/lt/veiklos-sritys/seima-ir-vaikai/socialine-parama-seimoms-ir-vaikams/pinigine-socialine-parama-nepasiturintiems-gyventojams-socialine-pasalpa-busto-sildymo-geriamojo-ir-karsto-vandens-islaidu-kompensacijos#soc.pasalpos_dydis (accessed on 15 June 2022). [38]

Ministry of Social Security and Labour (2017), *Įsakymas dėl užimtumo rėmimo priemonių įgyvendinimo sąlygų ir tvarkos aprašo patvirtinimo [Order on the approval of the conditions and procedure for the implementation of employment support measures]*, Republic of [25]

Lithuania.

OECD (2022), *Harnessing digitalisation in Public Employment Services to connect people with jobs*, OECD Publishing, Paris, https://oe.cd/digitalPES. [7]

OECD (2022), *Impact Evaluation of Vocational Training and Employment Subsidies for the Unemployed in Lithuania. Connecting People with Jobs.*, OECD Publishing, Paris, https://doi.org/10.1787/c22d68b3-en. [1]

OECD (2022), *Part-time employment rate* (indicator), https://doi.org/10.1787/f2ad596c-en (accessed on 5 April 2022). [14]

OECD (2022), *Reaching Out and Activating Inactive and Unemployed Persons in Bulgaria*, Connecting People with Jobs, OECD Publishing, Paris, https://doi.org/10.1787/7b91154a-en. [39]

OECD (2021), *OECD Tax-benefit model for Lithuania: Description of policy rules for 2020*, https://www.oecd.org/els/soc/TaxBEN-Lithuania-2020.pdf. [15]

OECD (2021), *Registered Unemployed and Job Vacancies dataset (database LAB_REG_VAC)*, OECD.Stat, https://stats.oecd.org/Index.aspx?DataSetCode=LAB_REG_VAC (accessed on 27 April 2021). [19]

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). [8]

OECD (2020), *Impact evaluation of labour market policies through the use of linked administrative data*, OECD, Paris, https://www.oecd.org/els/emp/Impact_evaluation_of_LMP.pdf. [9]

OECD (2020), *Impact Evaluations Framework for the Spanish Ministry of Labour and Social Economy and Ministry of Inclusion, Social Security and Migrations*, https://www.oecd.org/els/emp/Impact_Evaluations_Framework.pdf. [10]

OECD (2018), *OECD Reviews of Labour Market and Social Policies: Lithuania*, OECD Reviews of Labour Market and Social Policies, OECD Publishing, Paris, https://doi.org/10.1787/9789264189935-en. [17]

OECD (2002), *Glossary of Key Terms in Evaluation and Results Based Management*, OECD Publishing: Paris, https://www.oecd.org/dac/evaluation/2754804.pdf. [2]

Republic of Lithuania (2016), *Law on employment*. [24]

Sianesi, B. (2004), “An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s”, *Review of Economics and Statistics*, Vol. 86/1, pp. 133-155. [26]

Spector, S. (2022), “Should unemployment insurance cover partial unemployment?”, *IZA World of Labor*, https://doi.org/10.15185/izawol.199.v2. [40]

Užimtumo tarnyba (2022), *Aktyvios darbo rinkos politikos įgyvendinimo efektyvumas*, https://uzt.lt/darbo-rinka/aktyvios-darbo-rinkos-politikos-igyvendinimo-efektyvumas/ (accessed on June 2022). [6]

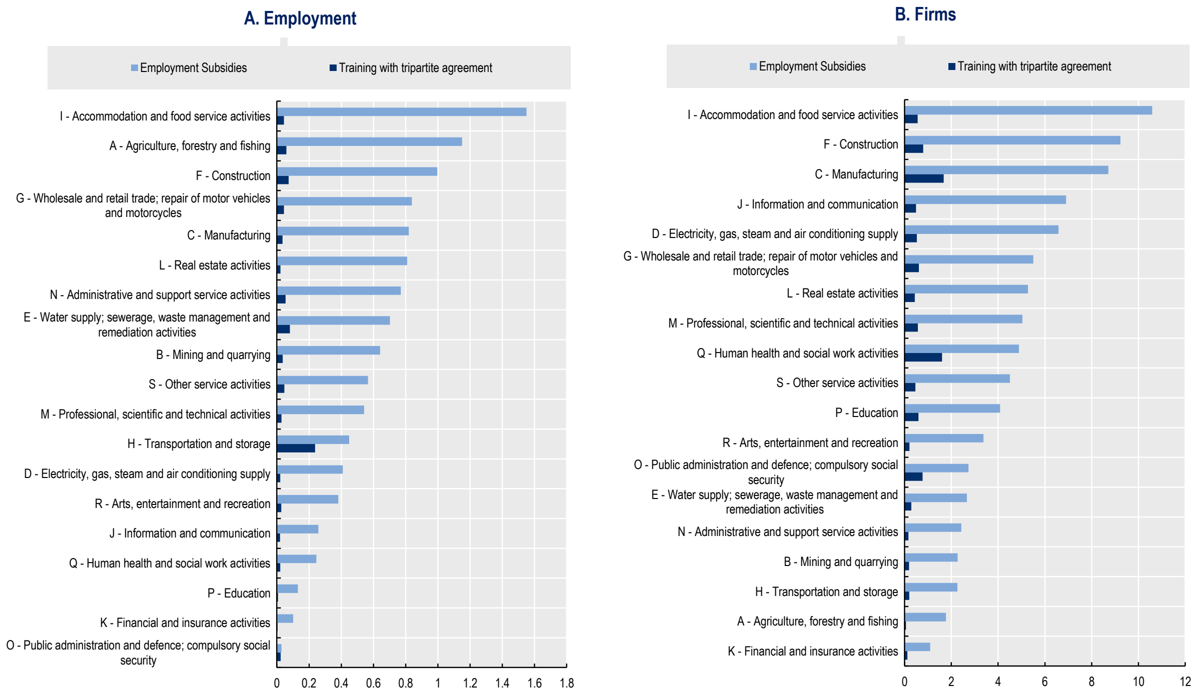
Užimtumo tarnyba prie Lietuvos Respublikos socialinės apsaugos ir darbo ministerijos (2020), *Dėl darbo rinkos stebėsenos vykdymo*, https://www.e- [5]

tar.lt/portal/lt/legalAct/f234e860f80511eaa12ad7c04a383ca0.

Annexe A. Additional figures

Figure A A.1. Incidence of vocational training with tripartite agreement and employment subsidies varies considerably across sectors of economic activity in Lithuania

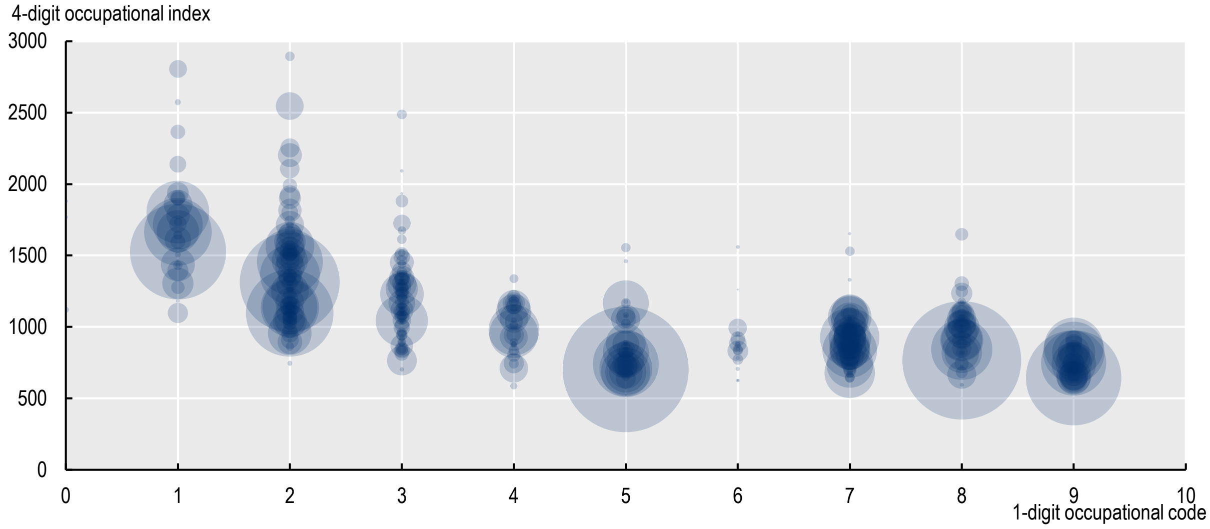
Share of employment and firms participating in ALMP within sector of economic activity



Note: Statistics refer to the 2018-20 period and do not include sole-proprietorships (which can also employ workers). Statistics on employment shares refer to annual averages during the 2018-20 period; statistics on share of firms relate to the share of firms having made use of the ALMP at any time during 2018-20.
Source: OECD calculations based on data from the Lithuanian Employment Service, Lithuanian State Social Insurance Fund Board and State Enterprise Centre of Registers data.

Figure A A.2. Occupational index is weakly related to 1-digit occupational code

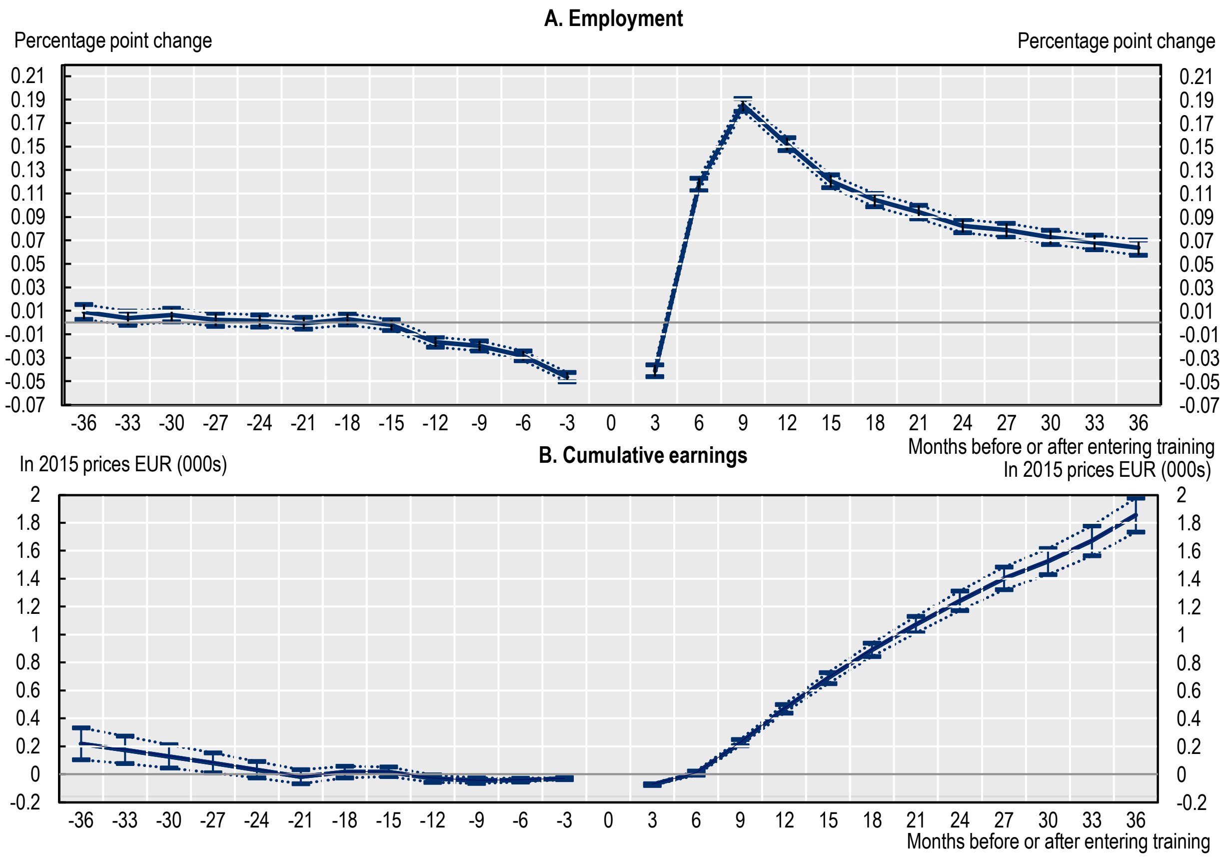
Relationship between 1-digit ISCO-08 codes and 4-digit occupational index



Note: Each shaded circle represents one 4-digit ISCO-08 code. The size of shaded circles is proportional to the number of individuals observed to be employed in the 4-digit occupation during the 2018-20 period.
Source: OECD calculations based on the Lithuanian Employment Service and Lithuanian State Social Insurance Fund Board.

Figure A A.3. Placebo tests show statistically insignificant effects for most time periods examined before individuals entered vocational training

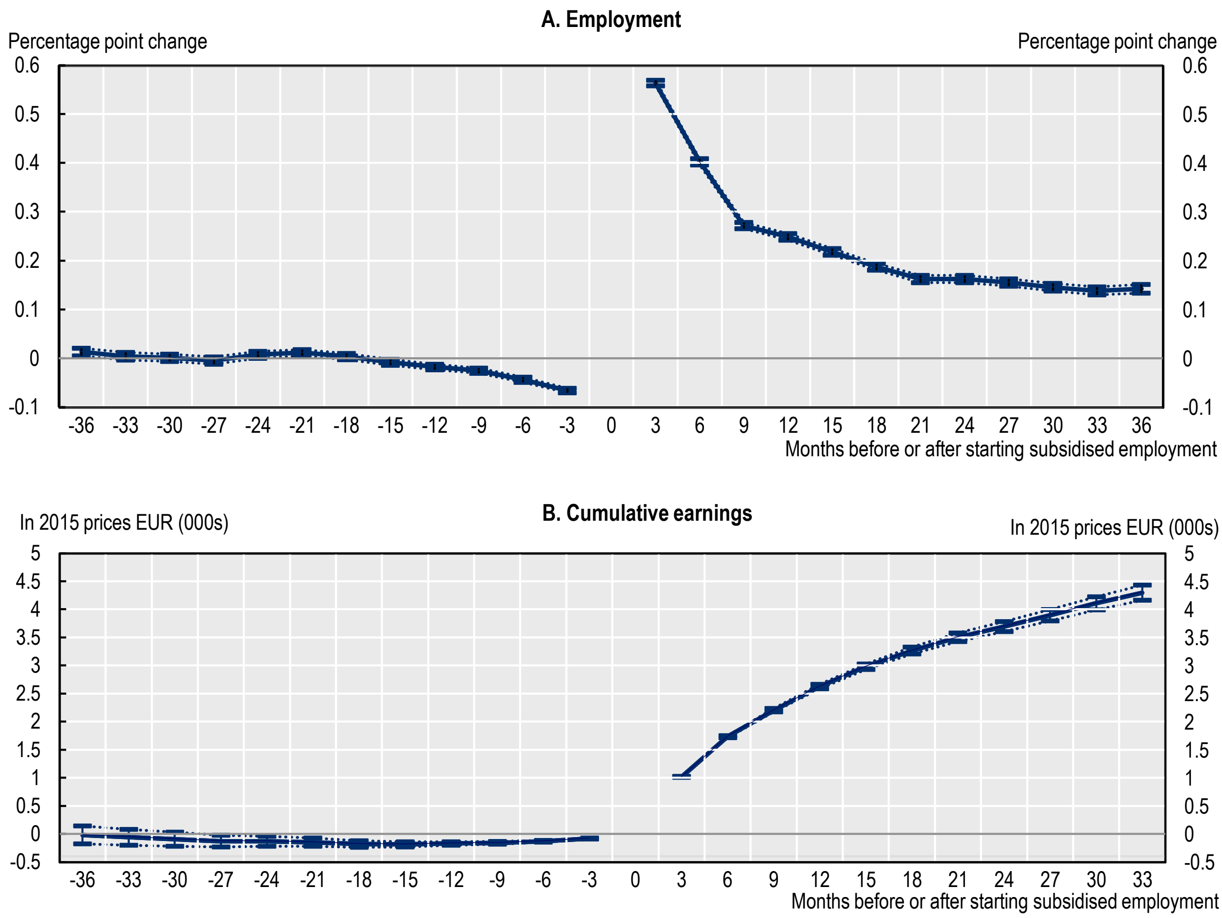
Percentage point change in employment probability (Panel A) and cumulative earnings (Panel B)



Note: The analysis presents nearest-neighbour propensity score matching results which matches individuals based on a number characteristics: duration of unemployment, age, gender, education, unemployment benefit receipt and level, language and other skills, municipality, employability and barriers to employment, as well as prior employment history and earnings. For every individual in the treatment group, the matching is conducted based on the values of these characteristics in the calendar month when the individual enters the programme. The comparison group is comprised of individuals with similar characteristics not entering ALMPs in that same calendar month. For paired individuals in the treatment and comparison groups, this calendar month is then the reference point after which outcomes are measured. The analysis is restricted to the region of common support. The standard errors are calculated based on the adjustment proposed by Abadie and Imbens (2016[36]). The confidence intervals are shown at the 5% level of significance and represented by the whiskers delimiting the dotted lines on the charts.
Source: OECD calculations based on the Lithuanian Employment Service and Lithuanian State Social Insurance Fund Board.

Figure A A.4. Placebo tests also show statistically insignificant effects for most time periods examined before individuals entered the employment subsidy programme

Percentage point change in employment probability (Panel A) and cumulative earnings (Panel B)



Note: The analysis presents nearest-neighbour propensity score matching results which matches individuals based on a number characteristics: duration of unemployment, age, gender, education, unemployment benefit receipt and level, language and other skills, municipality, employability and barriers to employment, as well as prior employment history and earnings. For every individual in the treatment group, the matching is conducted based on the values of these characteristics in the calendar month when the individual enters the programme. The comparison group is comprised of individuals with similar characteristics not entering ALMPs in that same calendar month. For paired individuals in the treatment and comparison groups, this calendar month is then the reference point after which outcomes are measured. The analysis is restricted to the region of common support. The standard errors are calculated based on the adjustment proposed by Abadie and Imbens (2016[36]). The confidence intervals are shown at the 5% level of significance and represented by the whiskers delimiting the dotted lines on the charts.
Source: OECD calculations based on the Lithuanian Employment Service and Lithuanian State Social Insurance Fund Board.

Annexe B. Additional tables

Annex Table 1. Data from Lithuanian Employment Service used in evaluation

List of attributes received by database

Unemployment registry database	ALMP database	Unemployment benefit database
Individual ID (pseudonymised)	Individual ID (pseudonymised)	Individual ID (pseudonymised)
Registration card ID (to identify individual unemployment spells)	Registration card ID (to identify individual unemployment spells)	Registration card ID (to identify individual unemployment spells)
Local office ID	ALMP name	Unemployment benefit Start date
Local office name	Vocational training programme type (over 2000 categories)	Unemployment benefit end date
Status	Vocational training tripartite agreement (yes/no)	Unemployment benefit Amount (initial)
Gender	Training type (formal/informal)	Unemployment benefit Amount after 3 months
Education level	Goal of training (e.g. certificate)	Unemployment benefit Amount after 6 months
Acquired qualification (education programme)	Training Start date	
Acquired qualification (occupation)	Training end date (planned)	
Date of qualification	Training end date (actual	
Date of Birth	Reason for termination	
Registration date	Certificate received	
Deregistration date	Subsidy amount in EUR (LTL before 2015)	
Reason for deregistration		
Municipality ID		
Municipality name		
Target occupation Code and Name (ISCO-4 digit)		
Work experience under occupation		
Target group		
Additional support criteria (for ALMP)		
English language skill level		
How many languages jobseekers speaks (including mother tongue)		
Driver's license		
Basic IT literacy		
Clerical work skills		
Social skills		
Motivation to work (low, medium, high)		

Annex Table 2. Data from SODRA used in evaluation

List of attributes received by database

Employment database	Earnings database
Panel A. Data on individuals who were registered as unemployed at some point during 2014-20	
Individual ID (pseudonymised)	Individual ID (pseudonymised)
Employer ID (pseudonymised)	Employer ID (pseudonymised)
Employment start date	Employment start date
Employment end date	Employment end date
Type of employment contract (e.g. permanent contract, apprenticeship; 9 different values)	Type of employment contract (e.g. permanent contract, apprenticeship; 9 different values)
Type of employment (description)	Type of employment (description)
Occupation (4-digit ISCO-08, 440 different values)	Earnings amount in EUR (LTL before 2015)
Panel B. Data on individuals who were never registered as unemployed during 2014-20	
Individual ID (pseudonymised)	Individual ID (pseudonymised)
Employer ID (pseudonymised)	Employer ID (pseudonymised)
Employment start date	Employer entity type (e.g. self-employed)
Employment end date	Month (year and month)
Type of employment contract (e.g. permanent contract, apprenticeship; 9 different values)	Earnings amount in EUR (LTL before 2015)
Type of employment (description)	Employment end date
Occupation (4-digit ISCO-08, 440 different values)	Type of employment contract (e.g. permanent contract, apprenticeship; 9 different values)
Gender	Type of employment (description)
Age group (5 year age categories)	Earnings amount in EUR (LTL before 2015)

Annex Table 3. Data from State Enterprise Centre of Registers used in evaluation

List of attributes received by database

Business registry database
Employer ID (pseudonymised)
Year
Employer institutional sector (39 categories)
Employer sector of economic activity (NACE level 1 - 22 categories)
Employer sector of economic activity (NACE level 4 - 872 categories)
Employer size (10 categories)
Employer earnings
Employer date from
Employer date until

Annex Table 4. Women and men undergo different types of training programmes

Top 10 most common vocational training programmes entered by men and women, based on presence of tripartite agreement with employer

Vocational training programme code and description	Number of entrants into programme
Panel A. Women without tripartite training agreements	
262081502- Manicurist training programme	1 513
361091301- Training programme for nurse assistants	1 388
262081106- Training programme for chefs	1 272
361072301- Training programme for nursing assistants	1 149
362081505- Training programme for hairdressers	1 033
262101201- Manicurist training programme	1 029
362041101- Training programme for accountants	1 025
362034401- Training programme for accountants	861
362081501- Training programme in decorative cosmetics	844
262101303- Chef training programme	808
Panel B. Women with tripartite training agreements	
362034401- Training programme for accountants	454
361072301- Training programme for nurse assistants	427
362041101- Training programme for accountants	176
262081106- Chef training programme	137
262054224- Training programme for tailor-operators	119
361076201- Training programme for assistant social worker	77
361091301- Training programme for nurse assistants	75
265084056- Training programme for drivers of motor vehicles for the transport of goods for the purpose of obtaining an initial vocational qualification on an urgent basis	65
265104108- Training programme for drivers of motor vehicles for the carriage of passengers for the purpose of obtaining an initial vocational qualification on an urgent basis	64
360072401- Training programme for dental assistants	61
Panel C. Men without tripartite training agreements	
265104112 - Training programme for drivers of motor vehicles for the carriage of goods for the purpose of obtaining an initial vocational qualification on an urgent basis	2 439
265084056- Training programme for drivers of motor vehicles for the transport of goods for the purpose of obtaining an initial vocational qualification on an urgent basis	2 057
262052104- Training programme for metal welder and electric and gas cutter	1 345
261084022- Training programme for forklift truck drivers	1 316
262058202- Finisher training programme	944
261104108- Forklift driver training programme	874
262058209- Plumber training programme	630
262071302- Training programme for an electrical installer for the installation and operation of electrical installations	619
262071504- Training programme for electric and gas metal welders and cutters	612
261062302- Training programme for woodcutters	594
Panel D. Men with tripartite training agreements	
265084056- Training programme for drivers of motor vehicles for the transport of goods for the purpose of obtaining an initial vocational qualification on an accelerated basis	3 231
265104112- Training programme for drivers of motor vehicles for the transport of goods for the purpose of obtaining an initial vocational qualification on an urgent basis	1 596
260084034- Recurrent vocational training programme for drivers of goods transport vehicles	497
265084052- Training programme for drivers of motor vehicles for the carriage of passengers for the purpose of initial vocational qualification on an urgent basis	467
223000487- Training programme for drivers of motor vehicles of categories C and CE	382
265104108- Training programme for drivers of motor vehicles for the carriage of passengers for	362

the purpose of obtaining an initial vocational qualification on an urgent basis	
260086203- Training programme for a worker in blast furnace work	307
223000549- Training programme for drivers of category C and CE vehicles	284
223000289- Non-formal vocational training programme for drivers of vehicles of categories C and CE	261
223000582- Training programme for drivers of vehicles of categories C and CE	223

Note: Programmes are ranked in terms of the total number of entrants during the 2014-20 period within panels; in the data, over two thousand programmes are observed in total. Programme descriptions are machine translations of the original descriptions in Lithuanian.
Source: OECD calculations based on data from the Lithuanian Employment Service.

Annex Table 5. Nearest neighbour matching by propensity score matching results in comparison group that is comparable to vocational training participants

Balancing properties for variables used in propensity score calculations (excluding municipality), vocational training participants

Variable	Mean: Treated	Mean: Non-treated (Unmatched)	Mean Control (Matched)	Difference in means: Treated vs. Non-treated	Difference in means: Treated vs. Control
Occupational index of last occupation (zero if not available)	699.832	657.860	699.130	41.972***	0.692
Occupational index of last occupation not available (1=n.a.)	0.208	0.262	0.208	-0.053***	-0.000
Log cumulative earnings (EUR in 2015 prices)	6.155	5.682	6.149	0.473***	0.007
Log cumulative days worked	4.322	4.000	4.319	0.322***	0.004
Cumulative days worked not available (1=n.a.)	0.245	0.299	0.245	-0.054***	-0.000
Cumulative earnings not available (1=n.a.)	0.245	0.299	0.245	-0.054***	-0.000
Age	38.359	43.714	38.339	-5.355***	0.021
Gender (1=women, 0=men)	0.469	0.488	0.470	-0.019***	-0.001
Unemployment benefit recipient (1=yes)	0.267	0.224	0.263	0.043***	0.004
Log level of unemployment benefits (zero if n.a.)	1.459	1.236	1.437	0.224***	0.023*
Drivers license? (1=yes)	0.591	0.484	0.587	0.107***	0.004
Education level: Primary	0.037	0.071	0.038	-0.034***	-0.001
Education level: Lower Secondary	0.181	0.167	0.183	0.014***	-0.002
Education level: Secondary	0.079	0.179	0.082	-0.101***	-0.003**
Education level: Vocational	0.000	0.000	0.000	0.000	0.000
Education level: Junior College	0.073	0.095	0.071	-0.022***	0.003**
Education level: University	0.090	0.116	0.088	-0.026***	0.002
Education level: Secondary With Vocational	0.486	0.321	0.484	0.165***	0.003
Education level: College	0.054	0.049	0.054	0.004***	-0.000
English language skills (1=yes)	0.252	0.177	0.250	0.075***	0.002
Basic IT skills (1=yes)	0.490	0.383	0.493	0.108***	-0.003
Clerical skills (1=yes)	0.069	0.062	0.070	0.008***	-0.001
Assessed motivation to become employed: (missing)	0.479	0.443	0.476	0.036***	0.002
Assessed motivation to become employed: Low	0.033	0.097	0.033	-0.064***	-0.000
Assessed motivation to become employed: Medium	0.339	0.319	0.341	0.020***	-0.002
Assessed motivation to become employed: High	0.150	0.142	0.150	0.008***	0.000
Emp. barrier: Persons up to 29 years of age	0.386	0.197	0.386	0.189***	0.000
Emp. barrier: Persons over 50 years of age	0.243	0.326	0.243	-0.083***	-0.000
Emp. barrier: Raising a child under the age of eight	0.110	0.071	0.112	0.039***	-0.001
Emp. barrier: Unskilled unemployed	0.054	0.067	0.054	-0.013***	0.000
Emp. barrier: Employed persons whose	0.000	0.000	0.000	-0.000	0.000

employers are on the COVID-19 list					
Emp. barrier: Working-age persons with a disability and a level of working capacity between 45–55	0.027	0.053	0.027	-0.026***	-0.000
Emp. barrier: Working-age persons with a disability and a level of working capacity between 30–40%	0.007	0.012	0.007	-0.006***	-0.000
Emp. barrier: Unemployed persons starting work after having acquired their qualifications	0.027	0.007	0.029	0.019***	-0.003***
Emp. barrier: Persons returning to the LM after being in prison	0.004	0.003	0.004	0.001***	-0.001**
Emp. barrier: Raising a disabled child under 18 years of age	0.003	0.002	0.003	0.000**	-0.000
Emp. barrier: Pregnant women	0.000	0.001	0.000	-0.000***	-0.000**
Emp. barrier: Persons who have been granted refugee status	0.000	0.000	0.000	0.000***	-0.000
Emp. barrier: Persons caring for family members	0.000	0.000	0.000	-0.000	-0.000

Note: Difference in means columns include significance tests denoting significance at the 10, 5, and 1 percent level respectively as *, ** and ***.
Source: OECD calculations based on data from the Lithuanian Employment Service.

Annex Table 6. Nearest neighbour matching by propensity score matching results in comparison group that is comparable to vocational training participants

Balancing properties for municipality binary variables used in propensity score calculations, vocational training participants

Municipality	Mean: Treated	Mean: Non-treated (Unmatched)	Mean Control (Matched)	Difference in means: Treated vs. Non-treated	Difference in means: Treated vs. Control
Alytaus miesto sav.	0.023	0.021	0.024	0.002***	-0.001
Druskininkų sav.	0.012	0.008	0.011	0.004***	0.001**
Alytaus rajono sav.	0.010	0.012	0.010	-0.002***	-0.000
Varėnos rajono sav.	0.009	0.010	0.009	-0.000	-0.000
Lazdijų rajono sav.	0.008	0.011	0.008	-0.003***	0.001
Birštono sav.	0.002	0.001	0.001	0.000***	0.000
Kauno miesto sav.	0.092	0.099	0.092	-0.007***	0.000
Jonavos rajono sav.	0.020	0.018	0.020	0.002***	0.000
Kaišiadorių rajono sav.	0.009	0.009	0.009	-0.000	-0.000
Kauno rajono sav.	0.028	0.028	0.029	0.001	-0.001
Prienų rajono sav.	0.009	0.008	0.010	0.001***	-0.000
Raseinių rajono sav.	0.011	0.012	0.011	-0.001***	-0.000
Klaipėdos miesto sav.	0.047	0.045	0.049	0.003***	-0.001
Neringos sav.	0.000	0.001	0.000	-0.000***	-0.000
Palangos miesto sav.	0.004	0.006	0.004	-0.001***	0.001
Klaipėdos rajono sav.	0.014	0.013	0.014	0.001	-0.000
Kretingos rajono sav.	0.007	0.008	0.007	-0.001***	-0.000
Skuodo rajono sav.	0.006	0.005	0.006	0.001**	0.000
Šilutės rajono sav.	0.021	0.018	0.022	0.003***	-0.001
Marijampolės sav.	0.018	0.019	0.018	-0.000	0.000
Vilkaviškio rajono sav.	0.017	0.015	0.016	0.001***	0.001
Šakių rajono sav.	0.011	0.010	0.010	0.000	0.000
Panevėžio miesto sav.	0.036	0.030	0.036	0.006***	0.000
Biržų rajono sav.	0.011	0.010	0.011	0.001***	-0.000
Kėdainių rajono sav.	0.012	0.016	0.012	-0.004***	-0.000
Kupiškio rajono sav.	0.010	0.008	0.010	0.002***	-0.000
Panevėžio rajono sav.	0.011	0.013	0.010	-0.002***	0.001
Pasvalio rajono sav.	0.013	0.011	0.014	0.002***	-0.001
Šiaulių miesto sav.	0.029	0.026	0.028	0.002***	0.000
Akmenės rajono sav.	0.011	0.010	0.011	0.001***	0.000
Joniškio rajono sav.	0.010	0.010	0.011	0.001*	-0.000
Kelmės rajono sav.	0.013	0.014	0.013	-0.000	-0.000
Pakruojo rajono sav.	0.010	0.008	0.010	0.002***	-0.000
Radviliškio rajono sav.	0.015	0.015	0.015	-0.001	-0.000
Šiaulių rajono sav.	0.012	0.015	0.012	-0.003***	-0.000
Tauragės rajono sav.	0.023	0.016	0.022	0.007***	0.001
Šilalės rajono sav.	0.010	0.007	0.010	0.004***	0.000
Jurbarko rajono sav.	0.017	0.014	0.016	0.002***	0.000
Mažeikių rajono sav.	0.027	0.024	0.027	0.004***	-0.000
Plungės rajono sav.	0.012	0.011	0.012	0.002***	-0.000
Telšių rajono sav.	0.015	0.014	0.015	0.002***	-0.000
Anykščių rajono sav.	0.009	0.011	0.009	-0.002***	-0.000
Zarasų rajono sav.	0.009	0.009	0.009	0.000	0.000
Ignalinos rajono sav.	0.009	0.009	0.009	-0.000	0.000

Molėtų rajono sav.	0.008	0.009	0.008	-0.000	-0.000
Rokiškio rajono sav.	0.017	0.014	0.017	0.003***	0.000
Utenos rajono sav.	0.015	0.015	0.015	0.000	0.000
Visagino sav.	0.011	0.011	0.012	0.001*	-0.001
Vilniaus miesto sav.	0.132	0.161	0.134	-0.029***	-0.003
Vilniaus rajono sav.	0.031	0.037	0.031	-0.005***	0.001
Trakų rajono sav.	0.008	0.009	0.008	-0.001*	0.000
Ukmergės rajono sav.	0.017	0.015	0.016	0.002***	0.001
Šalčininkų rajono sav.	0.015	0.015	0.013	-0.000	0.001**
Švenčionių rajono sav.	0.010	0.009	0.010	0.000	-0.000
Širvintų rajono sav.	0.006	0.006	0.007	0.001**	-0.000

Note: Difference in means columns include significance tests denoting significance at the 10, 5, and 1 percent level respectively as *, ** and ***.
Source: OECD calculations based on data from the Lithuanian Employment Service.

Annex Table 7. Nearest neighbour matching by propensity score matching results in comparison group that is comparable to employment subsidy participants

Balancing properties for variables used in propensity score calculations (excluding municipality), employment subsidy participants

Variable	Mean: Treated	Mean: Non-treated (Unmatched)	Mean Control (Matched)	Difference in means: Treated vs. Non-treated	Difference in means: Treated vs. Control
Occupational index of last occupation (zero if not available)	732.016	657.744	727.356	74.273***	4.641*
Occupational index of last occupation not available (1=n.a.)	0.182	0.262	0.185	-0.080***	-0.003
Log cumulative earnings (EUR in 2015 prices)	6.462	5.681	6.445	0.781***	0.018
Log cumulative days worked	4.546	3.999	4.535	0.547***	0.011
Cumulative days worked earnings not available (1=n.a.)	0.218	0.299	0.219	-0.081***	0.000
Cumulative earnings not available (1=n.a.)	0.218	0.299	0.219	-0.081***	0.000
Age	43.216	43.714	43.268	-0.498***	-0.048
Gender (1=women, 0=men)	0.466	0.488	0.474	-0.022***	-0.008**
Unemployment benefit recipient (1=yes)	0.263	0.224	0.265	0.040***	-0.002
Log level of unemployment benefits (zero if n.a.)	1.436	1.236	1.443	0.201***	-0.006
Drivers license? (1=yes)	0.560	0.484	0.553	0.076***	0.007**
Education level: Primary	0.052	0.071	0.053	-0.019***	-0.000
Education level: Lower Secondary	0.136	0.167	0.135	-0.031***	0.001
Education level: Secondary	0.165	0.179	0.164	-0.014***	0.001
Education level: Vocational	0.000	0.000	0.000	0.000	0.000
Education level: Junior College	0.109	0.095	0.110	0.014***	-0.002
Education level: University	0.123	0.116	0.122	0.007***	0.001
Education level: Secondary With Vocational	0.346	0.321	0.348	0.024***	-0.002
Education level: College	0.069	0.050	0.068	0.019***	0.001
English language skills (1=yes)	0.210	0.177	0.204	0.032***	0.006**
Basic IT skills (1=yes)	0.439	0.383	0.435	0.056***	0.004
Clerical skills (1=yes)	0.081	0.062	0.079	0.020***	0.003
Assessed motivation to become employed: (missing)	0.551	0.443	0.548	0.108***	0.003
Assessed motivation to become employed: Low	0.030	0.097	0.031	-0.067***	-0.001
Assessed motivation to become employed: Medium	0.291	0.319	0.291	-0.028***	-0.000
Assessed motivation to become employed: High	0.128	0.142	0.131	-0.013***	-0.002
Emp. barrier: Persons up to 29 years of age	0.338	0.197	0.342	0.141***	-0.003
Emp. barrier: Persons over 50 years of age	0.461	0.326	0.460	0.135***	0.001
Emp. barrier: Raising a child under the age of eight	0.092	0.071	0.093	0.021***	-0.002
Emp. barrier: Unskilled unemployed	0.058	0.067	0.056	-0.009***	0.002
Emp. barrier: Employed persons whose employers are on the COVID-19 list	0.000	0.000	0.000	0.000	0.000
Emp. barrier: Working-age persons with a disability and a level of working capacity	0.047	0.053	0.049	-0.005***	-0.002

between 45–55					
Emp. barrier: Working-age persons with a disability and a level of working capacity between 30–40%	0.027	0.012	0.028	0.014***	-0.001
Emp. barrier: Unemployed persons starting work after having acquired their qualifications	0.015	0.007	0.015	0.008***	0.001
Emp. barrier: Persons returning to the LM after being in prison	0.003	0.003	0.003	0.000	-0.000
Emp. barrier: Raising a disabled child under 18 years of age	0.002	0.002	0.002	-0.001**	-0.000
Emp. barrier: Pregnant women	0.000	0.001	0.000	-0.000***	0.000
Emp. barrier: Persons who have been granted refugee status	0.003	0.000	0.002	0.003***	0.000
Emp. barrier: Persons caring for family members	0.000	0.000	0.000	-0.000	0.000

Note: Difference in means columns include significance tests denoting significance at the 10, 5, and 1 percent level respectively as *, ** and ***.
Source: OECD calculations based on data from the Lithuanian Employment Service.

Annex Table 8. Nearest neighbour matching by propensity score matching results in comparison group that is comparable to employment subsidy participants

Balancing properties for municipality binary variables used in propensity score calculations, employment subsidy participants

Municipality	Mean: Treated	Mean: Non-treated (Unmatched)	Mean Control (Matched)	Difference in means: Treated vs. Non-treated	Difference in means: Treated vs. Control
Alytaus miesto sav.	0.033	0.021	0.032	0.011***	0.000
Druskininkų sav.	0.013	0.008	0.013	0.006***	0.001
Alytaus rajono sav.	0.013	0.012	0.013	0.001***	0.000
Varėnos rajono sav.	0.013	0.010	0.013	0.003***	0.000
Lazdijų rajono sav.	0.013	0.011	0.013	0.002***	0.000
Birštono sav.	0.002	0.001	0.002	0.001***	0.000
Kauno miesto sav.	0.072	0.099	0.071	-0.027***	0.001
Jonavos rajono sav.	0.019	0.018	0.020	0.002**	-0.000
Kaišiadorių rajono sav.	0.012	0.009	0.012	0.003***	0.000
Kauno rajono sav.	0.024	0.028	0.025	-0.004***	-0.001
Prienų rajono sav.	0.009	0.008	0.009	0.001*	-0.000
Raseinių rajono sav.	0.012	0.012	0.013	0.001	-0.001
Klaipėdos miesto sav.	0.045	0.045	0.042	0.000	0.003**
Neringos sav.	0.000	0.001	0.001	-0.000**	-0.000
Palangos miesto sav.	0.006	0.006	0.006	0.001**	-0.000
Klaipėdos rajono sav.	0.014	0.013	0.014	0.000	-0.000
Kretingos rajono sav.	0.008	0.008	0.007	-0.001	0.000
Skuodo rajono sav.	0.007	0.005	0.007	0.002***	0.000
Šilutės rajono sav.	0.020	0.018	0.019	0.003***	0.001
Marijampolės sav.	0.018	0.019	0.018	-0.001	-0.000
Vilkaviškio rajono sav.	0.016	0.015	0.016	0.001	-0.001
Šakių rajono sav.	0.012	0.010	0.012	0.002***	0.000
Panevėžio miesto sav.	0.032	0.030	0.032	0.002**	-0.001
Biržų rajono sav.	0.010	0.010	0.011	0.000	-0.001
Kėdainių rajono sav.	0.017	0.016	0.016	0.001*	0.001*
Kupiškio rajono sav.	0.009	0.008	0.010	0.001**	-0.000
Panevėžio rajono sav.	0.010	0.013	0.011	-0.003***	-0.000
Pasvalio rajono sav.	0.013	0.011	0.013	0.002***	-0.001
Šiaulių miesto sav.	0.032	0.026	0.032	0.006***	0.001
Akmenės rajono sav.	0.014	0.010	0.014	0.003***	-0.000
Joniškio rajono sav.	0.012	0.010	0.012	0.002***	-0.001
Kelmės rajono sav.	0.016	0.014	0.016	0.003***	0.000
Pakruojo rajono sav.	0.010	0.008	0.011	0.002***	-0.001
Radviliškio rajono sav.	0.018	0.015	0.018	0.002***	-0.001
Šiaulių rajono sav.	0.016	0.015	0.015	0.001**	0.001
Tauragės rajono sav.	0.017	0.016	0.017	0.001*	-0.000
Šilalės rajono sav.	0.008	0.007	0.008	0.001**	-0.000
Jurbarko rajono sav.	0.012	0.014	0.012	-0.002***	0.000
Mažeikių rajono sav.	0.029	0.024	0.030	0.005***	-0.001
Plungės rajono sav.	0.013	0.011	0.012	0.002***	0.000
Telšių rajono sav.	0.017	0.014	0.017	0.003***	0.000
Anykščių rajono sav.	0.011	0.011	0.012	-0.000	-0.001
Zarasų rajono sav.	0.011	0.009	0.011	0.002***	-0.000
Ignalinos rajono sav.	0.007	0.009	0.008	-0.002***	-0.001

Molėtų rajono sav.	0.008	0.009	0.009	-0.001	-0.001
Rokiškio rajono sav.	0.015	0.014	0.014	0.001***	0.001
Utenos rajono sav.	0.014	0.015	0.013	-0.001	0.000
Visagino sav.	0.012	0.011	0.013	0.001***	-0.001*
Vilniaus miesto sav.	0.118	0.161	0.119	-0.043***	-0.001
Vilniaus rajono sav.	0.032	0.037	0.032	-0.004***	0.000
Trakų rajono sav.	0.007	0.009	0.007	-0.001***	0.000
Ukmergės rajono sav.	0.015	0.015	0.014	0.000	0.001*
Šalčininkų rajono sav.	0.014	0.015	0.014	-0.001**	-0.001
Švenčionių rajono sav.	0.013	0.009	0.014	0.004***	-0.001
Širvintų rajono sav.	0.008	0.006	0.008	0.002***	-0.000

Note: Difference in means columns include significance tests denoting significance at the 10, 5, and 1 percent level respectively as *, ** and ***. Propensity score calculations and matching are done within individual calendar years.
Source: OECD calculations based on data from the Lithuanian Employment Service