Lucić and Puljiz *Journal for Labour Market Research* (2026) 60:13
https://doi.org/10.1186/s12651-026-00438-0

Journal for Labour Market Research

ORIGINAL ARTICLE

Open Access

"Make a wish": How a de-institutionalised care programme shaped the dynamics of women's aggregate labour market outcomes in croatia

Marko Lucić[1*] and Jakša Puljiz[2]

Abstract
Public works are usually considered a less important and less effective type of Active Labour Market Policies (ALMP) as has been emphasized by various empirical studies, mainly based on micro-level data. However, very few studies have looked at local-level labour market effects in the case of large-scale public works intervention. In this study, we perform a municipality-level counterfactual evaluation of a large-scale intervention targeting unemployed women without tertiary education in Croatia to work as assistants for elderly people. The municipality-level effects are examined in Croatia by applying difference-in-differences (DID) estimation and following the approach for staggered rollout by Callaway and Sant'Anna (Callaway and Sant'Anna 2021a). Empirical findings suggest that intervention has resulted in significant reductions in long-term unemployment over the longer term, albeit partly caused by the implementation dynamics, in particular the return of participants to the unemployment register after completion. More genuine and lasting impact on long-term unemployment was noted only two years after the municipality's exposure to the intervention. Despite the significant financial resources invested in the intervention, we find no evidence of displacement effects. We found no effects on women's employment in other jobs or on their activity rate. Besides the empirical findings, research also involves methodological insights regarding the application of aggregate-level counterfactual studies in the evaluation of labour market interventions, providing lessons on evaluation setup, evaluation timeliness, research questions and data availability.

Keywords Active labour market policies, Public works, Municipalities, Evaluation, Difference-in-differences

JEL Classification J08, J68, C23

1 Introduction

Active labour market policy (ALMP) generally refers to a set of measures aimed at enhancing the skills of unemployed individuals, assisting in job search efforts, subsidizing wages in both the private and public sectors, and overall improving labour market efficiency in terms of employability and earning capacity. These measures are particularly targeted toward vulnerable groups such as long-term unemployed, unemployed youth or low-skilled individuals (Kluve 2010; Martin 2015). In response to lessons learned from the 2008 financial crisis, the EU emphasized the risk of short-term unemployment

*Correspondence:
Marko Lucić
marko.lucic@idi.hr
[1]Institute for Social Research in Zagreb, Amruševa 11/II, Zagreb 10000, Croatia
[2]Institute for Development and International Relations, Ljudevita Farkaša Vukotinovića 2, Zagreb 10000, Croatia

evolving into persistent long-term unemployment, urging increased attention to this issue (Council of the European Union 2016). Before the COVID-19 pandemic, efforts to reduce long-term unemployment and activate inactive individuals had become integral components of ALMP. Various strategies were adopted to address complex employment barriers, including in-work coaching, counselling, follow-up, intra-agency cooperation, and public-private partnerships (Csillag 2021). However, ALMPs may also include demanding elements such as mandatory participation, monitoring, and financial sanctions, consistent with activation approaches that emphasise individual responsibility (Caswell et al. 2017). New programs were launched in several EU member states, with results demonstrating the effectiveness of personalized and intensive counselling combined with low caseloads. Additionally, a multidisciplinary approach, post-placement support, and a continuous learning process proved crucial (European Commission and ICON Institute Public Sector GmbH 2023).

Empirical studies confirm positive effects of active labour market policies (ALMPs) on the probability of securing employment (Card et al. 2018; Larsen 2002; Layard et al. 2005), However, these effects are typically modest, and they differ depending on the type of ALMP (Vooren et al. 2019). Despite the general trend of increasing support for ALMP across the EU over the past two decades, the content and financial significance of ALMP measures vary significantly among member states. These differences reflect national priorities and domestic legacies (Fargion and Profeti 2016). The most noticeable national variations pertain to the emphasis on employment compared to poverty reduction types of measures, where the latter focus more on material deprivation and access to social services.

One of the common measures included in ALMPs is public works. Public works or public sector employment is a type of measure aimed at increasing the demand for labour by directly employing those without a job in work programmes. Such programmes are usually implemented in the public or non-profit sector for a short amount of time and are based around community benefit (European Training Foundation 2022). They usually aim to reach dual objectives of providing temporary employment and generating some labour-intensive infrastructural projects and/or social services (Murgai et al. 2016; Subbarao et al. 2013). While they are quite important for some specific social goals, such as maintaining public infrastructure, especially in rural areas (European Training Foundation 2022), they are a less important type of intervention in terms of allocated resources (Robalino et al. 2013). Also, they are more common in low-income countries, where they play a vital role in income protection for the most vulnerable parts of the population (Ferreira and Robalino 2010; Malo 2018). In high-income countries, the rationale for such programmes is not limited to employment creation or inequality reduction. Public works are also justified by their potential to promote social inclusion and participation, particularly among groups whose long-term unemployment weakens labour market attachment and social integration. Empirical evidence shows that public works are particularly helpful for disadvantaged groups such as the long-term unemployed and older people, as well as populations in regions facing serious socio-economic challenges (Tzannatos 1999).

A substantial share of the empirical evidence on job creation schemes is based on micro-econometric evaluations focusing on individual participants' post-program employment probabilities. While these studies are essential for assessing participant-level outcomes, micro-level effects do not necessarily translate into aggregate labour market impacts. Large-scale interventions may generate general-equilibrium and composition mechanisms at the local level—such as displacement of regular jobs, reallocation of job search and hiring behaviour, and changes in unemployment register dynamics—that can attenuate, offset, or even reverse individual-level gains when observed at the municipality level. In this context, an aggregate counterfactual design becomes necessary to assess whether the intervention altered local labour market outcomes beyond the subsidised jobs it directly created.

In this paper, we use a counterfactual event study to evaluate the aggregate effects of a large-scale publicly sponsored employment (direct job creation) scheme on local labour markets. We evaluate implementation of a Croatian public employment programme called "Wish" in period 2018–2023. In Eurostat's Labour Market Policy (LMP) statistics for Croatia (Eurostat 2025), this programme falls under category 6 "Direct job creation" together with a measure named "Public works" (*javni radovi*). Both schemes provide temporary non-market jobs of community or social benefit that are additional to normal market demand and would not exist without public funding. Such schemes are typically implemented in the public or non-profit sector and are explicitly targeted at disadvantaged jobseekers (European Commission 2013; Subbarao et al. 2013). We therefore use the terms public works, direct job creation, and publicly sponsored employment interchangeably, while keeping Eurostat's categorisation as our empirical reference point.

While the programme targets low-skilled women in predominantly municipal (often rural) labour markets, its scale and design mirror a broader family of publicly sponsored employment schemes across Europe. Because the intervention reached a substantial share of the low-skilled female unemployed in most treated municipalities, its impacts speak to aggregate labour-market mechanisms that are transferable to other contexts where direct job creation is implemented locally. Because local units differ structurally in labour-market regimes and programme take-up, we focus on municipalities to preserve credible counterfactual trends (see Sect. 4). Despite the wide scope and the political and fiscal scale of the “Wish” intervention, evidence about its effects is relatively limited, which is unsurprising given that implementation lasted until the end of 2023.

We evaluate the programme’s aggregate labour market effects using three outcomes: (i) the share of long-term unemployed women within the target-group unemployment stock, (ii) women’s employment rate in jobs other than “Wish” (i.e., excluding subsidised programme employment), and (iii) women’s activity rate, all three at the municipal level. While the first outcome represents a key indicator of the programme’s effectiveness in reducing long-term unemployment, the other two outcomes provide complementary insights into broader labour market adjustments. First, to address concerns that exits to employment from the unemployment register may partly reflect mechanical programme entry rather than integration into regular jobs, we examine women’s employment outside the “Wish” scheme. Specifically, using administrative employment records, we construct an indicator capturing the share of employed women within the local population of working-age women, excluding women employed within “Wish” from the numerator. Second, we examine whether programme rollout affected women’s labour market activity at the municipal level. Conceptually, activity rates may respond if the intervention draws previously inactive women into registered unemployment prior to programme entry, or if programme participation strengthens labour market attachment beyond the subsidised employment spell. As activation is the primary proclaimed policy goal of public works and a usual benchmark of micro-level evaluations of public works, it is worth comparing aggregate, local-level effects with the micro-level effects in the literature. Beyond these outcomes, we explicitly assess whether the programme displaced regular employment opportunities for the target group. Finally, in addition to the empirical findings, the paper offers methodological insights into the design and interpretation of aggregate-level counterfactual evaluations of large-scale labour market interventions at the subnational level.

Building on the established evidence on job creation schemes, we formulate three research questions.

RQ1: Direct job creation may mechanically improve register-based indicators in the short run, but does it generate immediate or sustained reductions in long-term unemployment?

RQ2: If the programme strengthens labour market attachment among low-skilled women, does it contribute to increased activity rates and, over time, to more favourable unemployment dynamics?

RQ3: If the programme absorbs a sizeable fraction of low-skilled female labour supply, does the crowding-out of regular employment occur? However, in contexts with slack labour markets, this effect may be absent.

Our findings highlight that aggregate labour market responses to large-scale direct job creation schemes may differ across outcomes and time horizons. In particular, register-based unemployment indicators and broader employment-based measures need not move in tandem, underscoring the importance of evaluating multiple aggregate outcomes in parallel. The results also illustrate how implementation dynamics and institutional features of programme roll-out shape the interpretation of aggregate counterfactual estimates.

The remainder of this paper is structured as follows. The next section comprises a brief literature review. This is followed by a description of the intervention in Sect. 3, while Sect. 4 explains data used and the specifics of the intervention implementation and coverage at the LAU level to motivate the identification strategy in the following section. In Sect. 6, the results of the evaluation are presented, while in the final section, the local labour market impacts of the „Wish“ intervention are discussed, but also more generally, the assumptions, possibilities and limitations of the general identification strategy applied in the analysis.

2 Literature overview

Based on an analysis of policy responses to the 2008 crisis, Robalino et al. (Robalino et al. 2013) highlight the positive effects of direct public sector employment during times of crisis, particularly for informal-sector workers. Koltai (Koltai 2012) finds that public works schemes in Hungary had a positive impact on reducing inactivity and poverty levels between 2008 and 2010, demonstrating their buffering effects against the economic crisis. On the other hand, the literature generally recognizes the limited effectiveness of public works schemes in promoting long-term employment (Brown and Koettl 2015; Card et al. 2018; Kluve 2006; Kluve et al. 1999) and their inferior performance compared to other ALMPs (Kluve 2010; Martin and Grubb 2001; Pompili et al. 2023; Rodríguez-Planas and Jacob 2010).

Empirical studies largely confirm that continuous participation in public works schemes tends to decrease the likelihood of finding work in the primary job market. A meta-analysis of 57 micro-econometric evaluation studies by Vooren et al. (Vooren et al. 2019) shows that public job creation schemes have negative impacts on short-term employment probabilities but eventually become positive after 36 months. Several studies of job creation schemes in Germany document sizeable negative effects on regular employment (Hujer and Thomsen 2010; Thomsen and Walter 2010; Wolff and Stephan 2013). Brändle and Fervers (2021) show that even an innovatively designed job creation scheme failed to improve regular employment on average, yet negative effects were weakest among participants with very poor alternative employment prospects, implying that labour-market remoteness conditions effectiveness.

Country studies for the Central and Eastern European countries reach similar findings. Doválová and Hvozdíková (Dováľová and Hvozdíková 2014) analyze public works schemes in Slovakia, claiming that the chances of finding employment post-participation in public works schemes are lower than for a control group. Authors also identify various issues affecting post-participation employment outcomes, such as the absence of educational components in public works schemes and the inadequate capacity of public institutions to address the critical challenges participants face, such as skill deficits, health problems, and social exclusion issues. Cseres-Gergely and Molnár (Cseres-Gergely and Molnár 2015) analyzed public works schemes in Hungary between 2011 and 2013 and found that "…participating in public works for long hours and for a long time have a negative impact on the probability of entry to the open labour market and a positive impact on the probability of return to public works" (p. 156). Molnár et al. (Molnár et al. 2019) provide additional insight into the low effectiveness of public works schemes in Hungary, claiming that part of the problem is related to the decision to shift the task of tackling labour market problems largely from a competent network of employment offices to municipalities with no knowledge. In Croatia, evaluations of public works from 2009 to 2013 found a small positive impact on long-term employment post-participation (Croatian Employment Service and Ipsos 2016) and a small negative impact in the short term (Matković et al. 2012).

However, little is known about the aggregate impacts of public works, including their negative externalities. Beyond classic infrastructure-oriented public works, several European ALMPs have created additional jobs in care and household-related services, often targeting low-skilled women. Evaluations of such schemes (e.g., service voucher/domestic service programmes) provide a close benchmark for "Wish" and inform our expectations about lock-in risks and limited long-term integration.

Evidence from service voucher schemes suggests such programmes can generate sizable formal employment gains and shift work from the informal sector into regular employment relationships, though longer-term integration and job quality may remain concerns (Raz-Yurovich and Marx 2018; Adriaenssens et al. 2023). Relatedly, targeted wage subsidy and hiring credit evaluations show that employer-side instruments can support labour market entry for disadvantaged groups, but effects depend strongly on programme design and participant distance from the labour market (Jaenichen and Stephan 2011; Sjögren and Vikstrom, 2015).

To our knowledge, only a limited number of studies provide econometric counterfactual evidence on aggregate impacts of large-scale job creation or care-related employment schemes, including France, Ireland, Hungary and Belgium (L'Horty et al. 2024; OECD et al. 2024; OECD 2014; Szabó 2025; Leduc and Tojerow 2024). For example, the French TZCLD pilot reports reductions in long-term unemployment relative to synthetic controls (L'Horty et al. 2024), while Irish schemes are associated with declines in local long-term unemployment stocks (OECD et al. 2024). Evidence from Hungary suggests sizeable effects on registered long-term unemployment, but also negative spillovers on low-skilled wages (OECD 2014; Szabó 2025). A particularly relevant benchmark is Belgium's Service Voucher Programme, which targets low-skilled women through subsidised domestic services: Leduc and Tojerow (2024) find strong and lasting increases in female employment and reductions in unemployment and inactivity. While these results illustrate that care-related employment schemes can generate broader labour market effects, institutional design matters: unlike the market-based Belgian model, "Wish" primarily creates additional publicly financed care jobs, which may imply different spillover and displacement dynamics.

The EU Cohesion policy has been an important supporter of national labour market policies in member states through European Social Fund (ESF) funded programmes. This support has been particularly notable in the case of EU12 countries, where domestic funding for ALMP has been relatively low compared to older member states (European Commission 2010). In Croatia, most ESF funding during the 2014–2020 financial period was allocated to activities related to employment, education, and lifelong learning. Significant financial resources were invested in improving the employability of vulnerable groups, such as unemployed women, with the "Wish" intervention being the most prominent initiative. A qualitative interim evaluation of the ESF-funded measures in social inclusion conducted in 2021 concluded that the "Wish" intervention fared better in terms of social integration effects compared to other public works schemes (WYG 2021). However, the true effects of the "Wish" intervention on employability remained unexplored, which motivated this research. Additionally, this study aims to fill a gap in the literature, as it represents an addition to only a few evaluations known to us that assess aggregate impacts of direct job creation schemes, of which only two assess public works' negative externalities, as we do (in terms of crowding out employment in the open labour market).

The study also takes into account the fact that most direct job creation schemes are formally constrained by additionality and non-substitution rules—jobs should not compete with regular employment (Thomsen and Walter 2010). This institutional feature motivates our explicit aggregate test for displacement effects. Unlike the Belgian SVP program which subsidises household demand for domestic services (a market-based instrument), Wish creates additional publicly financed care jobs. This difference matters for expected displacement and medium-term integration of participants.

3 Description of the intervention

"Wish" is publicly sponsored employment whose target group are registered unemployed women with secondary school education at most. The programme has been explicitly designed as women-only. This design follows the official rationale stated in the programme documents: (i) low-skilled women are over-represented among the long-term unemployed in Croatia, and (ii) the programme simultaneously addresses unmet local demand for home-based elderly care. While this gender-specific targeting limits direct comparability with male jobseekers, it mirrors a broader European pattern of care-oriented ALMPs aimed at improving low-skilled female employment. The eligibility is not conditional on time spent in the register prior to becoming employed within the intervention. The participants provide care for the elderly in their local area with a wage close to the minimum wage. They work on employment contracts for the NGOs or the local government bodies that apply for the projects and administer them. The intervention's coverage is extensive – until the end of 2021, in 468 out of 556 local administrative units (LAUs), there was at least one participant by her place of residence. The calls for project proposals were three, whereby in the first call, participation of women could last up to 2 years, in the second call up to 18 months, and in the third call (with projects ongoing at the time of writing), participation lasts up to 6 months. The Phase 1 call was open from June 2017 to February 2020. The Phase 2 call was open from February 2020 to May 2021. Finally, the last call (Phase 3) was open from May 2022 to August 2022. Among the criteria for awarding projects in Phase 1 and Phase 2 were the county-level unemployment rate and one of the brackets of the Development Index at the LAU level. Importantly, the Phase 2 call coincided with the COVID-19 shock, which disproportionately affected elderly care demand and local labour markets; we therefore inspect dynamic effects by exposure time. However, time-frames of calls for proposals in different phases do not encompass the timing of entries to intervention for all participants, as project evaluation and selection, contract signing and kick-off took considerable time. This means that a substantial proportion of participants in a Phase began their work months after the close of a call.

A total of 7290 participants have participated in Phase 1, about 9100 in Phase 2 and about 8600 in Phase 3. Even at the national level, the intervention encompassed a significant share of its broad target group – for example, the number of entries into the intervention in 2018 made up about 6% of the average target group stock in the same year, i.e. all unemployed women with secondary education at most, while this target group covered almost half (46%) of the total annual average of all unemployed in Croatia. There were only 30 participants who entered the intervention in 2017, so the first quarter of 2018 can be taken as the first quarter of the intervention. Project implementation started in different quarters across LAUs, so the intervention setup is staggered treatment at the LAU level.

The ESF monitoring microdata collected for purposes of the so-called common longer-term result indicators (as of January 2024) offer some insight into the participants' post-program labour market status.: about 28% of participants were employed at any job at the point of six months after leaving the intervention and 16% were employed again within the "Wish" intervention at that time-point. Many women participated in the intervention more than once: out of 24,995 participants in total, there were 15,104 unique persons. For a somewhat flawed comparison, 21% of participants of public works financed from the ESF (other than "Wish") who exited the intervention from 2018 to 2022 were employed at the point of 6 months after leaving the intervention, at any job (subsidised or not). The serious flaw in this comparison is that we do not know the share of non-"Wish" public works participants who entered subsidised employment post-participation. We have only circumstantial evidence from the 2010–2013 period that around 30% participants of public works participated more than once (Croatian Employment Service and Ipsos 2016). Thus, the available data measured at the micro-level do not allow for an entirely valid comparison. However, even if we had entirely valid micro-level data, and even if we did a micro-level counterfactual study, that would not besufficient to conclude about the aggregate effects of the intervention at the level of local labour markets.

4 Data and sample

We analyze the impact of „Wish" on the labour market dynamics at the aggregate level. The units of observation are local administrative units (LAUs). The time periods observed are quarters. The ESF monitoring microdata on the intervention participants (2018–2023) were aggregated to the combination of the quarter of the participants' entry into the intervention and the participants' LAU of residence. From the ESF monitoring data, only these quarterly counts of participants' entries are used in the analysis. All the remaining data used comes from publicly available sources. Monthly data for each LAU on the the number of long-term unemployed within the target group (unemployed women with secondary education at most), the target group stock size and the number of target group entries to the unemployment register, were downloaded from the Croatian Employment Service (CES) interactive database "Statistika online" (Croatian Employment Service 2024) and aggregated to the quarterly level from 2012 to 2023. Monthly data for each LAU on the number of employed women were formed from the publicly available monthly statistical reports of the Croatian Pension Insurance Institute and aggregated to the quarterly level from 2013-Q3 to 2014-Q4. As the only age-by-gender population data in Croatia at the LAU-level come from population censuses (taken in 2011 and 2021), a proxy was used for time-varying LAU-level counts of working-age women, which is a denominator needed to calculate the second and the third outcome (women's employment rate in other jobs and women's activity rate). To approximate the LAU-level counts of working-age women, the LAU-level shares of working-age women among all women were taken from the 2021 population census and applied to the yearly LAU-level population estimates for women. The comparison of the differences between the shares of working-age women between the 2011 and 2021 population censuses across LAUs did not point to systematic differences.

Women's employment in other jobs was calculated as a ratio of the quarterly mean of the employed women and the estimate of working age female population in the municipality, whereby counts of "Wish" participants were previously subtracted from the monthly counts of employed women. Women's activity rate was calculated taking the sum of the quarterly means of the employed and the unemployed women as the numerator, whereby "Wish" participants were included in the monthly counts of employed women.

Other LAU-level data used as covariates in the pre-treatment period were included in the calculation of the Development Index (Ministry of Regional Development and EU Funds 2018) and originally collected from the Croatian Bureau of Statistics, the CES and the Ministry of Finance: 2011 population size, population difference between 2011 and 2021, average general unemployment rate 2014–2016, developmental group of LAUs according to the 2013 development index, 2011 share of the population aged 16–65 with at least secondary school and the average municipality original revenue 2014–2016.

We operationalized the treatment intensity as follows: we divided the quarterly aggregated number of resident participant entries into the intervention (from the ESF microdata) by the quarterly average stock of the residing target group in the unemployment register (from the CES data). We took the peak of this ratio within an LAU over the observed period of entries (2018–2023) as the treatment intensity. As the quarterly sums of entries were divided by the quarterly average target group stocks, this ratio could be larger than 100, so we top-coded the treatment intensity to 100. We took the first quarter with any number of entries into the intervention as the quarter when the treatment began for a municipality.

Figure 1 shows the treatment intensity across LAUs. Treatment intensity allocation is dependent upon explicit design features of the intervention (project selection criteria, which included the county-level unemployment rate and the LAU-level Development index) and the local demand, the latter perhaps coupled with political pressure (having in mind that many of the projects were administered by the local government). Both groups of factors have to do with the differences in the labour market position of the target group, including the availability of lower-qualified jobs related to the tourist season, whereby the intervention was almost completely absent among residents of Istria and of most of the Dubrovnik-Neretva County, two coastal and highly developed tourist regions. There may be other, less obvious factors of the local treatment intensity, such as the lack of capacity to administer projects or closeness to LAUs that were more agile in applying for and administering projects which might explain the absence of treatment or its very low intensity in some municipalities where (a more intensive) treatment would be expected (e.g. the light coloured areas in the inland parts of Adriatic Croatia in Fig. 1).

Croatia has two types of LAUs: 428 municipalities and 128 towns. In general, the labour market position of the target group is better in urban than in rural parts (which also applies to jobseekers in general), which is the very likely explanation of the fact that the treatment in municipalities was on average almost twice more intensive at its peak than in towns (33% vs. 18%). This may confound the relationship between treatment and the outcomes, besides other unobserved (economic, social and demographic) differences between towns and municipalities. For this reason, we restricted the analysis to municipalities.

Only the 369 municipalities whose at least one resident participated in the intervention were kept in the analysis. The rationale for eliminating other units, where there were no participants at all, is that they are likely more different from the units with the most intensive treatment than the units with at least some, albeit a very small number of participants, since self-selection of the units without participants into treatment was completely absent

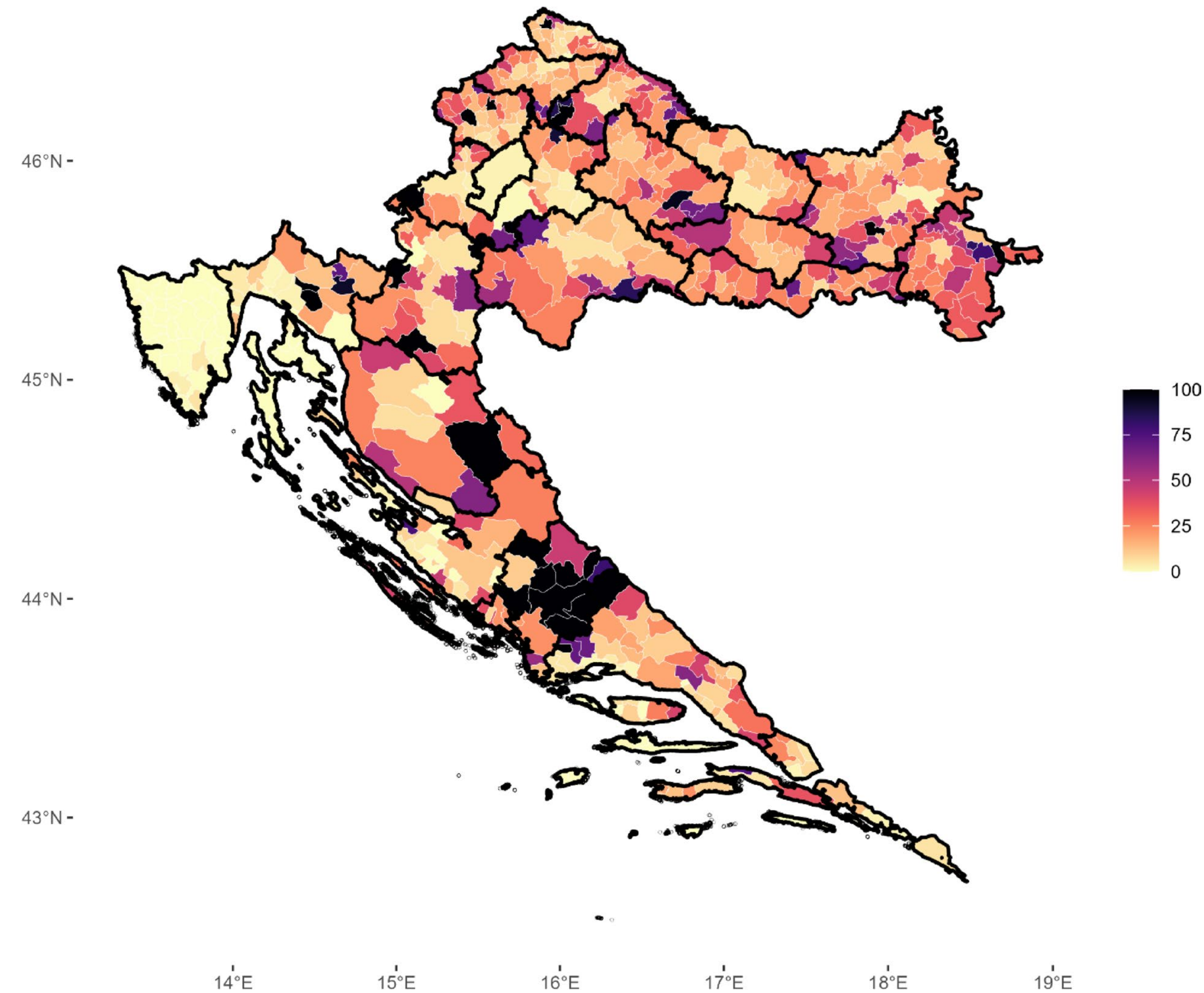

Fig. 1 Peak quarter treatment intensity in the local administrative units (2018–2023). Colours are scaled according to the highest quarterly ratio of the sum of entries into the intervention and the target group stock (3-month average number of registered unemployed women without tertiary education) within a LAU in 2018–2023. For the 25 LAUs with this ratio above 100, the treatment intensity was top-coded into 100. The black lines represent county (NUTS-2) borders

and thus those would provide a less comparable control group.

An important feature of the intervention set-up are discontinued entries of participants over time within each municipality as the three open calls were published in 2017, 2020 and 2022 ("Phase 1", "Phase 2" and "Phase 3"). Very rarely, there were multiple projects within the same call and the same municipality, which resulted from limited local pools of eligible women and probably also from the pooling of the resources needed to administer the projects within each municipality. The median period between the first two top quarters ranked by treatment intensity within each municipality was 9 quarters, which is mainly a consequence of Phase 1 and Phase 2 projects' starting times within municipalities. For illustration, Fig. 2 shows treatment intensity in each of the 10% most intensively treated municipalities until the end of 2023.

The differences in the patterns of the peaks are related to the different time patterns of the local projects funded through the "Wish" intervention. Figure 2 also demonstrates that in some case municipalities were participating in all three calls (three major peaks), but there were also cases of participation in two or only one call, as has been the case with Žumberak, last municipality shown in the figure.

The next figure shows the trends in three outcomes that we analyze as dependent variables Trends in those outcomes serve as the first indication of the effectiveness of the intervention on the targeted group at the local level, without taking differences in municipalities' characteristics and intervention starting times into account. The figure shows the means of the three outcomes in deciles of treatment intensity encompassing the bottom 40% and the top 20% of municipalities by treatment

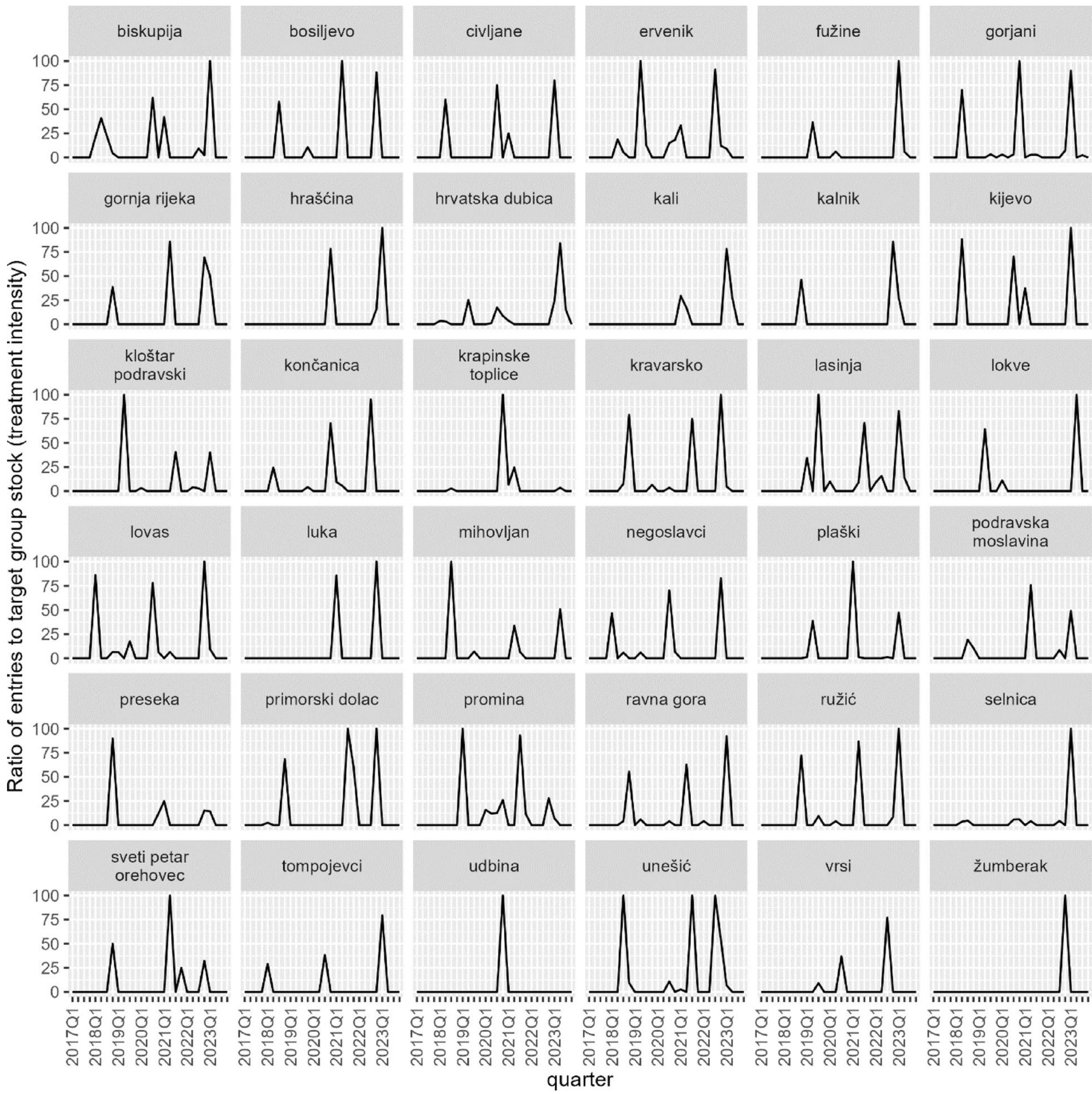

Fig. 2 Treatment intensity over time (quarterly ratio of the sum of entries into the intervention and the target group stock) for the 10% most intensively treated municipalities (2018–2023)

intensity, whereby decile 10 refers to the highest level of treatment intensity. The vertical dashed grey line in Q1 2018 represents the beginning of the intervention, i.e. the first municipalities entering treatment. The series were smoothed by four-quarter-wide moving averages.

The general trend of decreasing the share of long-term unemployed in the target group before the intervention started (top panel in Fig. 3) is likely attributable to the economic recovery after the Great Recession (prolonged in Croatia). The occurrence and magnitude of the post-treatment COVID-19 shock to the share of LTU women within the target-group unemployment stock (growing in 2020–2021), women's employment rate in other jobs (decelerating or stagnating during the pandemic) and women's activity rate (seemingly unchanged) seem to be independent of treatment intensity, if existent, which is encouraging from the standpoint of the identification strategy employed in this paper (difference-in-differences).

Descriptives of the municipalities' characteristics by deciles of treatment intensity at its peak quarter in each municipality are shown in Table 1. The pre-intervention characteristics shown in Table 1 were selected among all those available at the municipality level based on

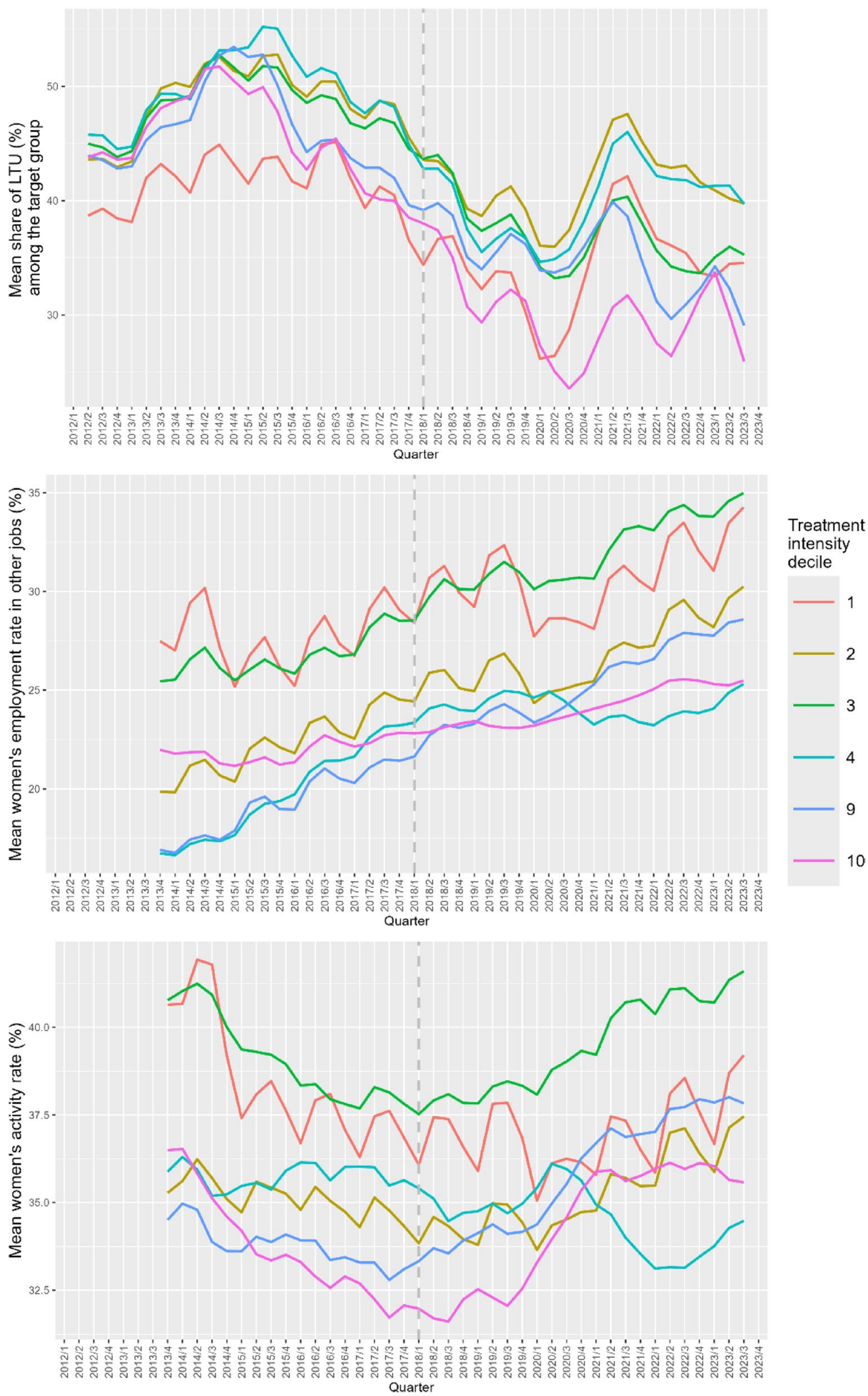

Fig. 3 Mean share of long-term unemployed women within the target-group unemployment stock, women's employment rate in other jobs and women's activity rate over the observed period within the bottom four and the top two deciles of peak treatment intensity (smoothed by four-quarter-wide moving averages)

Table 1 Municipalities' target group stock and pre-intervention characteristics by deciles of peak quarter treatment intensity

Deciles of peak quarter treatment intensity	*n*	Min-max peak quarter treat. intensity	Mean peak quarter treat. intensity (%)	Mean target group stock at peak quarter treat. Intensity	Mean population 2011	Mean rel. pop. change 2011–2021 (%)	Mean share of educated pop. 16–65 yrs (2011)	Mean unemployment rate 2014–2016	Mean Development Index 2013	Mean municipality revenue per capita 2014–2016 (standardized)
1	37	1.1–7.1	4.7	52.7	4,187	-6.3	73.9	14.8	0.92	104.48
2	37	7.4–11.1	9.5	58.3	3,940	-9.8	70.2	17.7	0.81	98.94
3	37	11.3–17.0	14.2	55.4	3,813	-12.9	68.0	17.4	0.77	98.20
4	37	17.3–21.0	19.0	67.3	3,251	-15.4	62.6	22.8	0.67	95.57
5	37	21.0–26.2	23.5	79.1	3,604	-19.5	63.0	25.0	0.62	93.58
6	37	26.3–31.9	28.9	41.1	2,662	-15.6	62.1	20.4	0.69	96.71
7	37	32.0–37.9	34.9	50.6	2,957	-17.3	62.5	22.3	0.65	95.39
8	37	38.0–47.9	42.6	40.3	2,617	-15.7	63.3	20.1	0.68	96.35
9	37	48.0–75.0	60.9	30.0	2,112	-18.3	60.5	21.7	0.68	98.01
10	36	75.9–100.0	94.6	15.9	1,788	-19.0	61.9	18.7	0.68	97.36

theoretical considerations (i.e. their importance for the local labour market dynamics) which was confirmed by their significance in the pooled ordinary least squares models (not shown). The least and the most intensively treated municipalities differ considerably on all of the selected pre-intervention characteristics relevant to the local dynamics of employment. Compared to the least intensively treated municipalities, the most intensively treated municipalities had on average smaller populations in 2011, larger negative population change until 2021 (mostly due to emigration), smaller shares of population with at least secondary school in 2011 and larger general unemployment rates 2014–2016 (except for unemployment rate in the fourth decile). This is expected as the disproportionately high coverage of areas with more severe developmental difficulties was intended by the intervention design – among the criteria for awarding projects were the county-level unemployment rate and one of the 5 brackets of the Development Index at the LAU level[1]. LAU characteristics shown in Table 1 are measured at the LAU level and used as covariates in subsequent inferential analyses to control for between-LAU differences in those characteristics when estimating the impacts of the intervention.

5 Methodology

The identification method used is difference-in-differences (DID) estimation of the period-specific average treatment effects on the treated (ATTs) for staggered rollout (Callaway and Sant'Anna 2021a) that lends itself naturally because the first pick-up of the intervention was not simultaneous in all municipalities. The treatment and control groups in the DID analyses are various groupings according to peak treatment intensity i.e. according to the target group inflow vs. stock ratio in the quarter when this ratio was the highest within a LAU (top 20 vs. bottom 40%, top 20 vs. bottom 20% and bottom 20 vs. top 10%). Treatment entry for each LAU is the first quarter with any recorded participants' entries by her place of residence.

Though we begin with a continuous treatment intensity variable (the peak quarterly ratio of the number of entries and the target group stock within each municipality), we chose to discretize the treatment variable into treatment and control groups of municipalities after considering the strict assumptions of continuous treatment in DID (Callaway et al. 2024). Namely, continuous treatment implies further assumptions of unbiased estimators in addition to the DID assumptions in the standard case of binary treatment group indicator: (1) effect linearity, meaning that the percent change of the dependent variable

[1] The 2013 Development Index was among the criteria for Phase 1 and the 2017 Index for Phases 2 and 3.

with the unit change of treatment intensity is identical across the entire range of treatment intensity; (2) conditional parallel trends assumption applies to different levels of treatment intensity; (3) there is no endogenous self-selection into treatment intensity. These assumptions are highly unrealistic in the case of our setup. In particular, as regards assumption (1), the more intensive the treatment (i.e. the larger the share of the target group that participates), the larger the "offset" in the dependent variables that can be expected upon municipalities' entry into treatment, especially in the case of exits to employment from the register - here this is expected almost by definition since the participants' entry into the intervention implies exiting to employment from the register. As regards assumptions (2) and (3), the endogenous self-selection into various treatment intensity levels is built into the design and pick-up of the intervention, as the intervention was intended to focus on areas with more labour market difficulties (hence the criteria of project selection mentioned in Sect. 3 and the applicants met with higher political pressure and/or demand for starting projects at the local level where the labour market difficulties were more pronounced. For these reasons, we divided the sample into deciles of treatment intensity and repeated the analyses for three different subsamples i.e. three different definitions of the treatment and the control group of municipalities. These repeated analyses constitute one of the results' robustness checks. Among the three definitions used (bottom 40% vs. top 20%, bottom 20% vs. top 20% and bottom 20% vs. top 10% by treatment intensity as the control and treatment group, respectively), the first is our preferred definition for a few reasons. This definition bears the largest statistical power (sample size). It encompasses the LAUs in the 3rd and 4th decile of treatment intensity (designated as control) that are more similar to the top deciles (designated as treated) in pre-treatment characteristics, as can be seen in Table 1 above. It also encompasses more intensively treated units in the control group than the other two definitions and thus makes for a more conservative (downwardly biased) estimate of the effect.

The DID estimators of interest are the average treatment effects on the treated by the length of the municipalities' exposure to the treatment. Let G_i denote the quarter in which municipality i first enters the intervention, with $G_i = 0$ for never-treated municipalities. In the first step, we estimate group–time average treatment effects $ATT_{g,t}$ for each treated cohort $g > 0$ and each post-treatment quarter $t \geq g$, comparing municipalities first treated in quarter g to never-treated municipalities, conditional on time-invariant municipality characteristics (described in six rightmost columns of Table 1).

In the second step, we aggregate the estimated group–time effects into dynamic event-time effects ATT_e, where event time is defined as $e = t - g$. These dynamic effects summarize the average impact of the intervention e quarters relative to first treatment and are reported in an event-study format. The weights w_g are proportional to cohort size, ensuring that larger treatment cohorts receive greater influence in the aggregated estimates. We chose the ATT_e estimators among all estimators described by Callaway and Sant'Anna (Callaway and Sant'Anna 2021a) because they offer insight into the average treatment effect dynamics regardless of the different intervention starting times for different municipalities. Henceforth, we call these estimators "the DID event study estimators" for brevity reasons[2].

$$ATT_{g,t} = E\left(Y_{it}(1) - Y_{it}(0) \middle| G_i = g\right), \quad t \geq g$$

$$ATT_e = \sum_g \left(w_g \cdot \mathbf{1}(g + e \leq T) \cdot ATT_{g,g+e}\right)$$

$$w_g = \frac{N_g}{\left(\sum_{h>0} N_h\right)}$$

As a placebo randomization test of significant estimates, we randomly reassign treatment status and entry quarters across municipalities within the estimation sample 499 times, preserving both the number of treated municipalities and the empirical distribution of treatment entry quarters. We then re-estimate the event-study specification and compare the observed post-treatment effects to the resulting permutation distribution. As a robustness check, we conduct a global randomization-inference test based on the maximum absolute post-treatment t-statistic from the event-study, and evaluate its significance using the permutation distribution generated by random reassignment of treated municipalities and entry quarters.

$$S = max_{e \in E_{post}} \mid ATT_e \, / \, se(ATT_e) \mid$$

The final step in defining the estimation samples was the exclusion of the 11 municipalities in the two upper deciles of treatment intensity that formed too small treatment entry cohorts for the DID estimates to be reliable – from Q3 2019 onwards (treatment entry cohorts of sizes from 1 to 3 LAUs). A complete overview of the formation of the estimation samples is in Table 2.

The resulting rollout structure of the estimation sample in the preferred specification (top-20 vs. bottom 40% of treatment intensity) is entirely situated at the earlier periods as quarters of entry, from 2018 Q1 to 2019 Q2 (the

[2] These estimators are analogous to the event study estimators in the more often DID setup of single-period treatment, with an important difference that all different periods (in our case, quarters) of entry into treatment are coded as 0.

Table 2 The formation of the estimation samples

Step	Estimation samples
1.	From the initial 556 LAUs, 128 towns were excluded and only municipalities were retained.
2.	Of the 428 municipalities, 369 with at least one participant by her place of residence were retained.
3.	Among the 73 municipalities from the 9th and 10th treatment intensity deciles, 11 were excluded due to their treatment entry cohorts being too small.
Final	Three estimation samples are three definitions of the control and treatment groups according to treatment intensity: bottom 40% vs. top 20% (148 control LAUs vs. 62 treated LAUs), bottom 20% vs. top 20% (74 control LAUs vs. 62 treated LAUs) and bottom 20% vs. top 10% (74 control LAUs vs. 28 treated LAUs).

frequencies of municipalities entering treatment being 10, 13, 9, 14, 10 and 6, respectively). This diminishes the importance of the usual caveat of staggered treatment DID – weaker support for the later post-entry periods.

All estimates in the next section are estimated with the *did* package in the R statistical environment (Callaway and Sant'Anna 2021b; Core Team 2024). First-step regression estimators were used to estimate the parameters, and their standard errors were computed using the multiplier bootstrap procedure, clustering on municipalities.

6 Results

6.1 The share of long-term unemployed women within the target-group unemployment stock

The DID event study estimates of the intervention's impact on the share of LTU within the target group are shown in Fig. 4 for three different treatment and control group definitions (rows), with and without covariates (columns). Significant impacts are those whose confidence intervals (green vertical lines) do not cross the horizontal dashed line (zero). Looking at the upper two specifications (i.e. two upper rows of Fig. 4), significant impacts lowering the share of LTU occurred in three periods and rather late, at the 9th quarter into the intervention, at the 15th – 16th quarter into the intervention and in the three last quarters observed (21st – 23rd quarter since the intervention onset). The size of the first two impacts is around a 9% points difference in the share of LTU, and the size of the third impact is more than double that of the previous two, around 22% points lower share of LTU in the treated municipalities.

These impacts are partly spurious, arising from the dynamics of the intervention implementation, which is revealed by introducing another dependent variable – the number of entries into the unemployment register – estimating the intervention's impacts on this outcome too (bottom row of Fig. 5) and comparing the latter to the impacts on the share of LTU (repeated at the top row of Fig. 5 from the top row of Fig. 4). Namely, the positive estimates of the impact on entries of the target group into the unemployment register coincide with the ones when the negative impact on the share of LTU was found (bottom row of Fig. 5). In other words, significantly lower shares of LTU in the treated municipalities can be partially ascribed to the change in the denominator, i.e. to an increase of the target group size due to participants' return to the unemployment register after completing a stint within the intervention. To put things in the context of the intervention setup more precisely, participation lasted 8 quarters in Phase 1 projects, 6 quarters in Phase 2 projects which followed Phase 1 immediately in most municipalities and 2 quarters in Phase 3 which started 4 quarters after Phase 2 stopped accepting new projects (see Fig. 2 for illustration). This timeline corresponds to the 8th, 14th and 22nd quarters after the intervention onset within an average municipality, and these are precisely the quarters in which we found positive impacts on the entries into the unemployment register and which immediately precede the quarters with significant differences in LTU share. However, the impact on LTU share in the last three observed quarters does not entirely fit into this explanation, as the significant impact on LTU share is first found in the 21st quarter and lasts through the next two quarters, while the impact on entries into the register is found only in the 22nd quarter (albeit not significant at the 5% level). The latter leaves room for the interpretation that the long-term impact of the intervention (5 years and onwards since the intervention onset) was genuine i.e. not predominantly arising from the change in the denominator.

The placebo random assignment in the top-20 versus bottom-40% treatment intensity estimation sample confirms the baseline results, both in the case of the global randomization-inference test (panel A of Fig. 6) and in the case of individual event times (panel B of Fig. 6).

The LAU level of development, measured by the official Development index, was one of the criteria for project selection in line with the intervention's aim to help women in more disadvantaged local labour markets. Heterogeneity analysis of the impact on the share of LTU women within the target-group unemployment stock, shown in the top row of Fig. 7, suggests that the intervention did not diminish the share of LTU women within the first two years in the more developed municipalities (panel A1). In the later periods, from the 20th quarter into the intervention onwards, the point estimates in the less and the more developed subsamples show the same pattern of increasing impact with similar magnitude, the last quarter in the more developed subsample notwithstanding, though wider confidence intervals due to smaller subsamples relative to the baseline models in Fig. 4 deem the impacts insignificant in the more developed municipalities. The bottom row of Fig. 7 points to larger

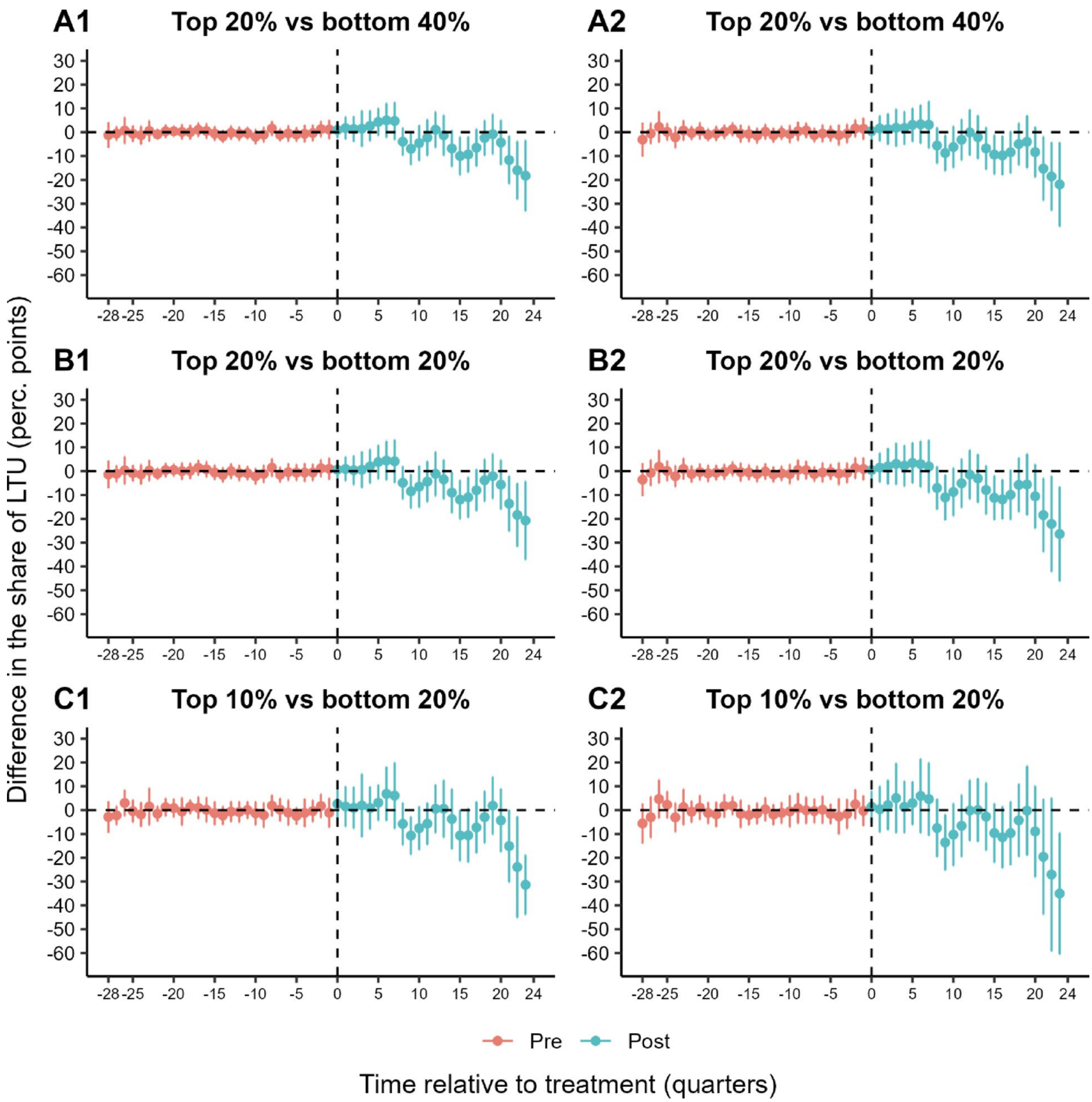

Fig. 4 The DID event study ATT estimates for three different definitions of the treatment and control groups with 95% confidence intervals: the share of long-term unemployed women within the target-group unemployment stock (left column: without covariates; right column: with covariates). Covariates included in models shown in the right column are LAU population size in 2011, relative population change 2011–2021, the share of educated people in the 16–65 age group, mean unemployment rate 2014–2016, development index in 2013 and mean municipality revenue per capita 2014–2016

effects for municipalities with a higher ratio of 60 + year-olds versus 0–19 year-olds in the latest intervention periods

6.2 Women's employment rate in other jobs

Figure 8 reports the DID event study ATT estimates for women's employment outside "Wish". Across treatment-control definitions and model specifications, the estimated effects are statistically insignificant throughout the observed time period. Importantly, we do not observe negative deviations between programme phases that would indicate crowding-out of regular jobs. At the same time, the absence of positive effects suggests that subsidised employment created by the programme did not translate into measurable aggregate gains in women's regular employment at the municipality level.

Taken together, these results imply that "Wish" operated primarily as an additional publicly financed employment channel, without detectable

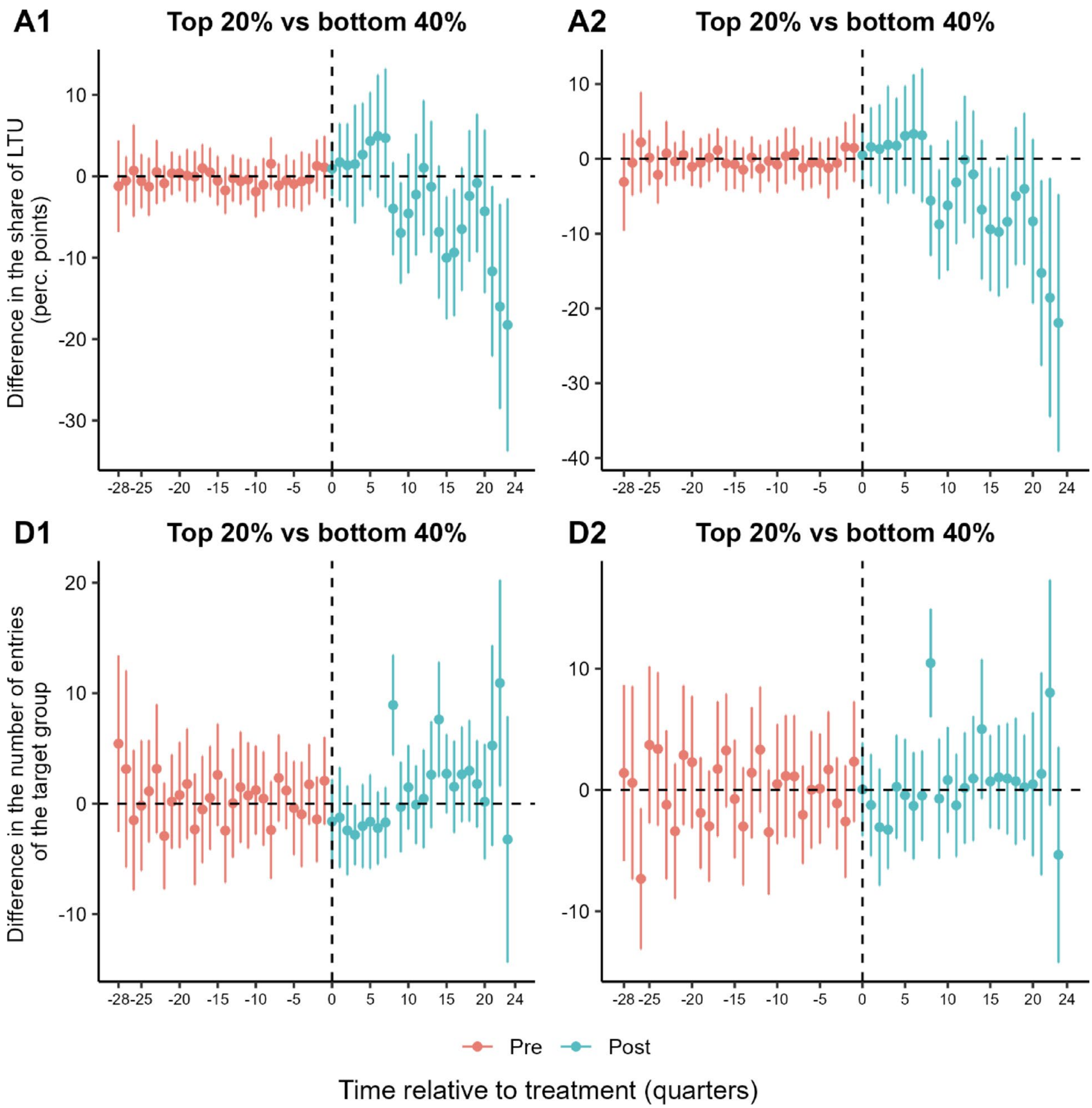

Fig. 5 The DID event study ATT estimates with 95% confidence intervals: the share of long-term unemployed and entries into the unemployment register within the target group (left column: without covariates; right column: with covariates). Covariates included in models shown in the right column are LAU population size in 2011, relative population change 2011–2021, the share of educated people in the 16–65 age group, mean unemployment rate 2014–2016, development index in 2013 and mean municipality revenue per capita 2014–2016

spillovers—positive or negative—on women's employment outside the programme.

6.3 Women's activity rate

The DID event study ATT estimates for women's activity rate are reported in Fig. 9, using the same three alternative definitions of treatment and control groups as in the previous subsections, and presenting specifications both without and with covariates. Across all definitions and model specifications, the estimates remain statistically insignificant. In other words, we do not find robust evidence that more intensive exposure to the "Wish" programme altered the local activity rate of women in treated municipalities relative to control municipalities during the observed post-treatment period. Nevertheless, the pattern of the point estimates is informative. The estimated coefficients are mostly positive, which could suggest a small activation effect; however, given that confidence intervals consistently include zero, we refrain from interpreting this pattern as a reliable programme

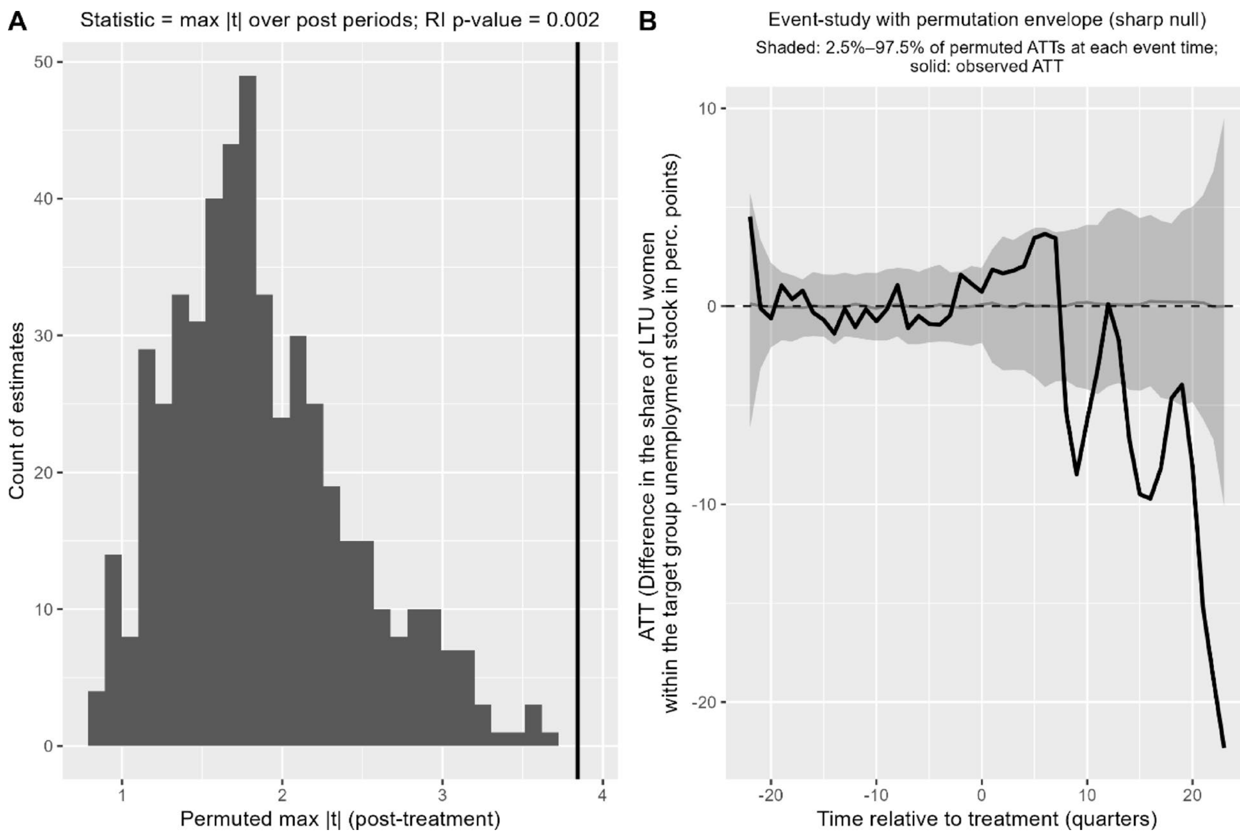

Fig. 6 Placebo random assignment tests of the impact on the share of LTU women within the target-group unemployment stock: global randomization inference test (**A**) and the comparison of permuted ATTs with the observed ATT (**B**). Covariates included in the models are the same as in the models shown in Fig. 4

impact. This lack of statistical significance may reflect several factors, including limited statistical power for detecting small activation effects at the municipality level, the gradual nature of activation processes, and the fact that the programme is primarily designed around employing registered unemployed women rather than directly targeting inactive women.

7 Discussion and conclusion

This study examined the labour market impacts of the "Wish" intervention, a large-scale public works programme targeting unemployed women without tertiary education in Croatia. Employing a difference-in-differences event study design, we analysed aggregate-level administrative data from municipalities to assess the effects on long-term unemployment and local employment. Our findings can be summarised in three policy-relevant aggregate outcomes. First, we detect statistically significant reductions in the share of long-term unemployed women within the target-group unemployment stock, yet these effects partly reflect implementation dynamics, particularly increased inflows into the unemployment register after programme completion. The co-evolution of impacts on the share of LTU women in the target-group unemployment stock and on the entries into the unemployment register, as well as the findings that Phase 2 projects, mostly initiated in 2021, yielded reductions in the LTU share comparable to Phase 1 projects that started before the pandemic, suggest that implementation dynamics, rather than COVID, shaped the development of the former. This is corroborated also by a more stable pattern of significant effects in the last observed quarters, when COVID-related restrictions had been lifted. Second, we find no evidence that the intervention affected women's employment outside the subsidised jobs created within "Wish". This implies that the programme neither generated aggregate "step-stone" effects into regular employment nor displaced regular employment opportunities at the municipal level. This pattern is consistent with the broader micro-evidence on job creation schemes, which frequently reports lock-in effects during participation and limited impacts on regular employment outcomes (Card et al. 2018; Pompili et al. 2023). Our aggregate evidence complements this literature by showing that even when a job creation scheme is implemented at scale, the measurable spillovers into non-programme employment may remain limited. Questions may be posed as to what stands behind non-existent

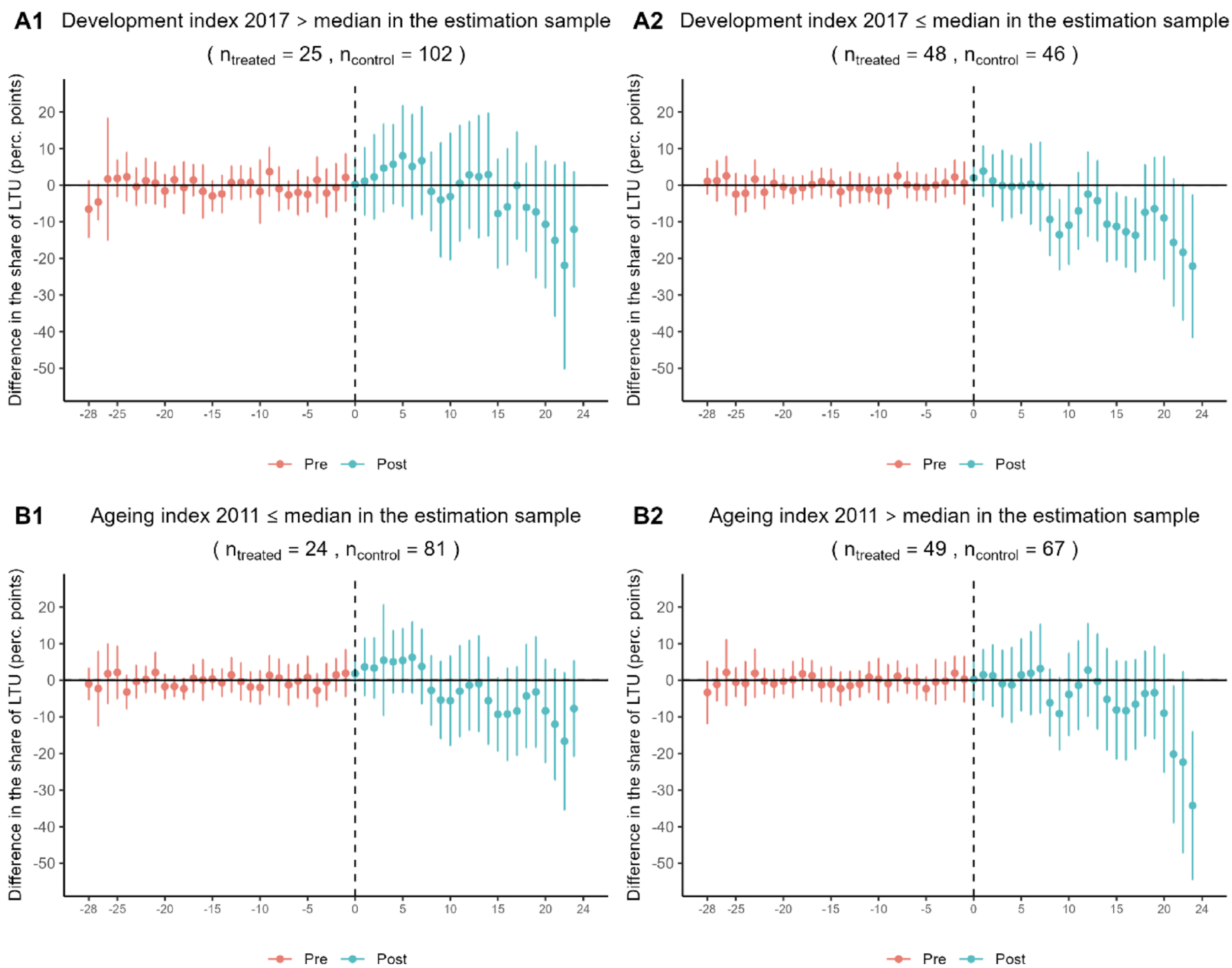

Fig. 7 Heterogeneity of impact on the share of LTU women within the target-group unemployment stock by municipalities' Development index 2017 and Ageing index 2011 within the estimation sample (top 20% vs. bottom 40% treatment intensity) with 95% confidence intervals. Covariates included in the models are the same as in the models shown in Fig. 4

crowding out effects on employment at the aggregate level. Cross-sectoral displacement effects hidden in the aggregate statistics (e.g., a flow of workers from tourism and hospitality to the programme) are not a likely occurrence, due to a combination of the programme's low remuneration and the scarcity of tourism-related, lower-qualified jobs in the most intensely participating municipalities, counted among the treated groups.

Third, we do not detect robust effects on women's activity rates, suggesting that the programme did not substantially alter aggregate labour market participation during the observed period. This is in contrast with the micro-level effects of public works found in the counterfactual study for the 2010–2013 period in Croatia (Croatian Employment Service and Ipsos 2016), which were significant and relatively large (8–11% points). One of the possible explanations for this discrepancy could be that 2010–2013 was a period of deep recession and the expansion of ALMP in Croatia, which might have increased the public works participants' post-participation attachment to the unemployment register in comparison to other unemployed who did not participate in any ALMP at the time. Another possible explanation would take into account that the participants of "Wish" are all women and emphasize the weaker female attachment to the labour market (Nestić and Tomić 2018), which could have been manifested in pre- and post-participation inactivity of the participants. One possible mechanism explaining the absence of an effect on additional jobs outside the programme is that the programme mainly formalised care work that had previously been performed informally, thereby bringing it into official employment statistics for the duration of participation. However, given the nature of informal work, testing this possibility would require dedicated data collection. The absence of an effect on women's activity rates provides only indirect evidence against this interpretation: if the programme had mainly drawn previously inactive women

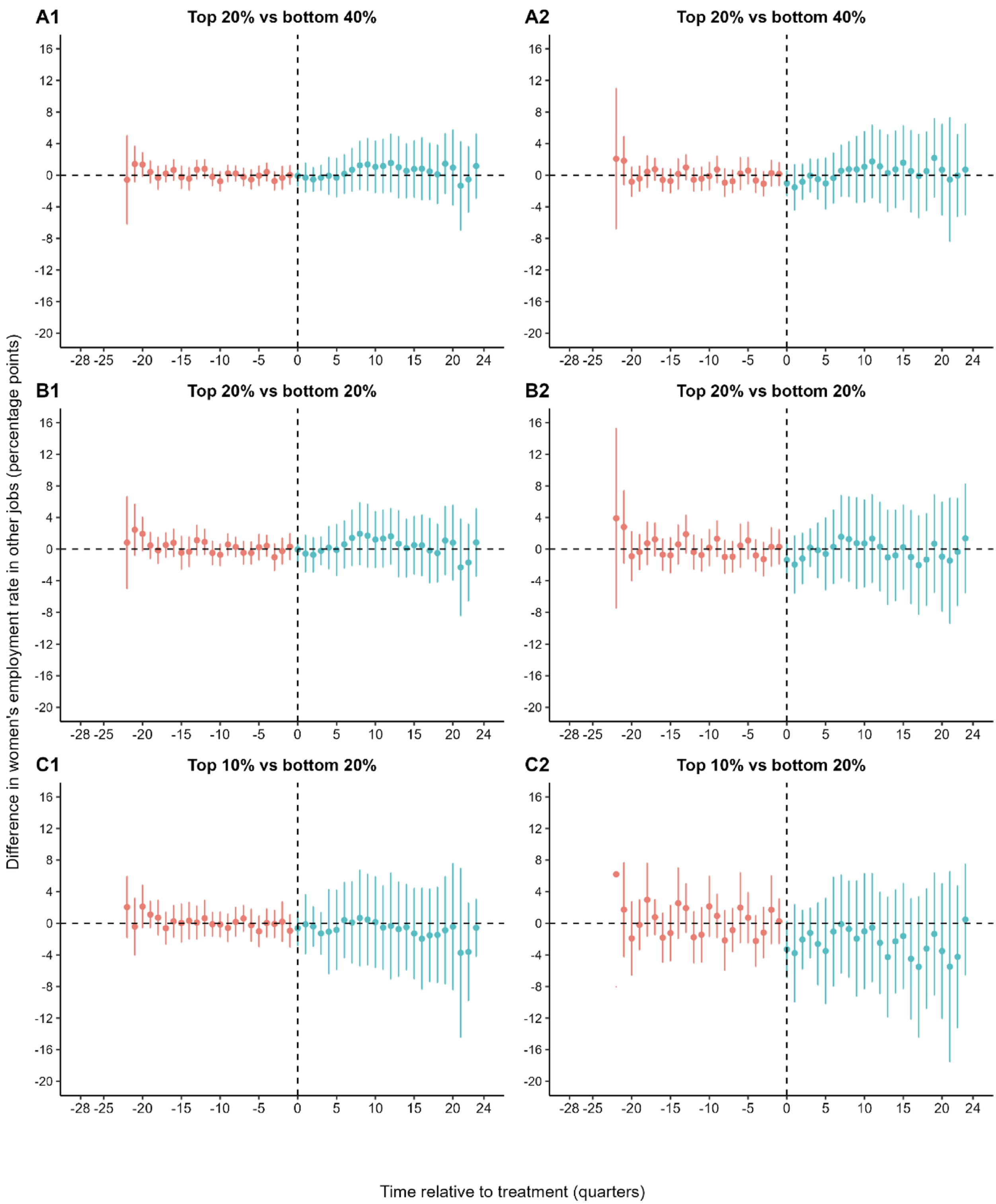

Fig. 8 The DID event study ATT estimates with 95% confidence intervals: women’s employment rate in other jobs (left column: without covariates; right column: with covariates). Covariates included in models shown in the right column are LAU population size in 2011, relative population change 2011–2021, the share of educated people in the 16–65 age group, mean unemployment rate 2014–2016, development index in 2013 and mean municipality revenue per capita 2014–2016

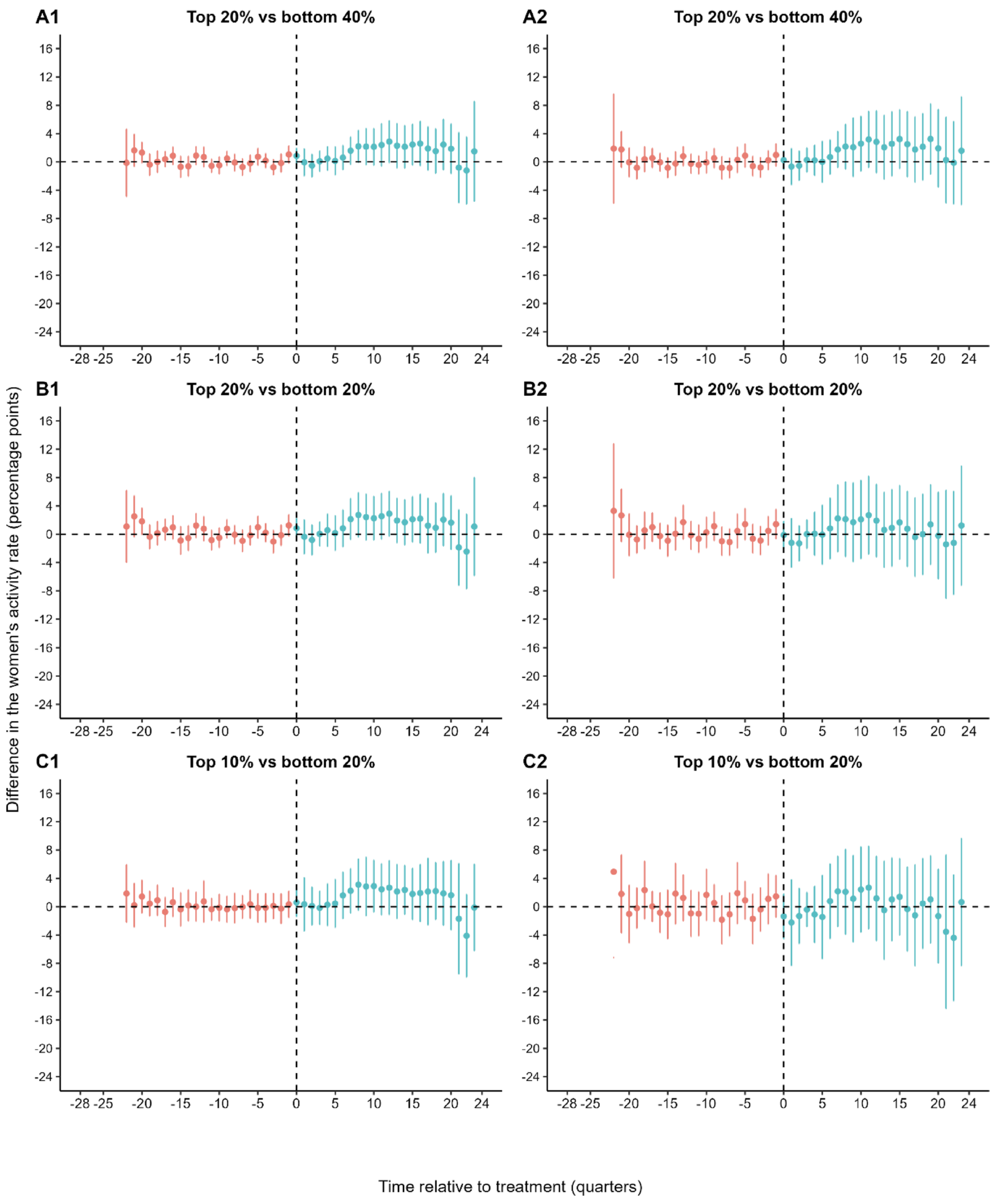

Fig. 9 The DID event study ATT estimates with 95% confidence intervals: women's activity rate (left column: without covariates; right column: with covariates). Covariates included in models shown in the right column are LAU population size in 2011, relative population change 2011–2021, the share of educated people in the 16–65 age group, mean unemployment rate 2014–2016, development index in 2013 and mean municipality revenue per capita 2014–2016

into formal employment, it should also have increased women's activity rates. While it is likely that this was happening to a considerable extent, it is less likely that such activation was a strong or dominant aggregate effect of the programme.

From a cost-benefit perspective, the annual average unit cost of 4,462 euros per participant in the 2018–2022 period corresponds to around 18,500 euros spent per one LTU person less in the unemployment register[3]. We did not find directly comparable estimates of cost-benefit in terms of reducing LTU within the unemployment register for any of the ALMPs. However, we can contextualize our estimate in other indicators. In Croatia, public works in 2013 yielded 1 person retained in activity for approximately 22,500 euros (Croatian Employment Service and Ipsos 2016). The domestic services subsidy in Belgium created one job for 13,512 euros in 2013 (Leduc and Tojerow 2024).

Beyond the empirical findings, this paper offers methodological insights regarding the application of aggregate-level counterfactual studies in the evaluation of labour market interventions. Several key lessons emerge regarding evaluation setup, evaluation timeliness, research questions and data availability.

Regarding the evaluation setup, for a research question about impacts at the aggregated level to make sense, the scale of the evaluated intervention at the observed level has to be relatively large, which usually also means relatively wide target groups with extensive coverage by the intervention as is the case here. Within this general case, depending on the intervention evaluated, sometimes the research questions at the aggregate level will be far more adequate than the micro-level ones or perhaps even the only type of evaluable questions due to the intervention setup which makes finding a valid comparison group of individuals (including the geographical criterion and all other strong labour market barriers and enablers) very difficult. In the case of „Wish", which covered more than 20% of women without tertiary education within a municipality at the median treatment intensity (see Table 1), micro-level evaluations, regardless of the method used (DID, matching, etc.) are difficult if not impossible to do correctly.

Another important prerequisite is knowledge about the policy mix for the target group addressed by the evaluated intervention. Rarely will there be only one policy addressing the target group's labour market status. Not taking this into account may obscure the impact of the focal intervention. If the differences in coverage by other policies between the observed units are stable over time, or, what is closer to reality, the coverage by other interventions can be controlled for in the analysis, it is possible to proceed. Often, there is a good chance that the data about participation in other relevant interventions will be available at the aggregated level and not at the micro-level.

The number of treated units at the observed i.e. aggregate level does not have to be large. Synthetic control methods, recently extended for the staggered roll-out (Ben-Michael et al. 2022; Porreca 2022), offer a way to assess the impact for a small number of treated units, even only one, which is an advantage over micro-level studies. The DID analysis which takes into account statistical uncertainty (i.e. beyond the simplest 2×2 case) demands more units, but not overwhelmingly so. Even the DID extension for staggered roll-out applied here, the group-time ATTs (Callaway and Sant'Anna 2021 a) estimated for each combination of the period and the treatment entry cohort (i.e., the length of exposure), can function with as little as 5 units having the same length of exposure[4]. In our analysis, the units are geographical (LAUs), but depending on the intervention, units may also be economic sectors or age groups – for example, Tomić and Žilić, performed a DID evaluation of traineeships in Croatia, which is based on comparing age groups (Tomić and Zilic 2020).

Regarding evaluation timeliness and the related concerns about research questions, impact evaluations are often regarded as necessarily long-term endeavours whose results need to be waited for a long time, which is then emphasized as the main flaw of such evaluations from the perspective of policy-making. We argue that the latter perspective assumes evaluating the effects on post-participation outcomes of individuals, while aggregate-level analyses such as ours enable evaluating the intervention's early impact on the target group's labour market position at the level of the observed units if the regularly updated aggregated data are used.

Moreover, the usual distinction between process and impact evaluations, while reserving the counterfactual methodology for the latter, does not need to be applied. Counterfactual event studies at the aggregate level may offer early insights into the added value of the intervention, both in terms of outputs and results. In the intervention analyzed here, the outputs would be exits to employment of the target group as the intervention consists of the employment of the target group, while results are viewed as the participants' labour market position in the post-intervention period and may be operationalized

[3] Taking the impact in the last observed quarters (around 20% points) and the difference in peak treatment intensity between the control and the treated group (83% points), it is (4,462 euros * 83)/20 = 18,517 euros.

[4] 9 The *did* R package warns about counts less than 5, but there are no exact lower boundaries on the number of units here. The assessment of how many units are enough depends on the context, e.g. the existence of unit-specific shocks during the post-treatment period or other differences between the currently treated and other units.

as the number of long-term unemployed in the target group after introducing the intervention. If the early added value in terms of outputs is of primary interest and the evaluated intervention is temporally concentrated (as in the case of „Wish"), the contribution of other factors affecting situation of the target group could be less present, which can be enabling for conclusions about the true impact of the intervention. Early counterfactual evaluations of results, not only outputs at the aggregate level may inform policy-making too. As a hypothetical example, if an evaluation of the „Wish" intervention's impact on long-term unemployment like this one was done within the first year of implementation, it might have been an argument for targeting the long-term unemployed within the original target group, as it would have shown early on that there was no effect on the share of long-term unemployed women within the first year.

As a final point on the possibilities of the approach used here regarding research questions, event studies such as this one may be used to identify displacement effects on employment at the aggregate level, which the micro-level studies generally cannot (Escudero 2018).

Regarding data availability, as illustrated here, the analyses of the large-scale interventions' impacts on the local labour market dynamics in Croatia can be done predominantly using publicly available aggregate data, mostly thanks to the CES interactive database *Statistika online*, which offers a wide range of job-seekers' characteristics to filter. However, relying on the unemployment register only for data on outcomes is both limiting and potentially harmful for the validity of any before-after research design, as the unemployment register data do not necessarily reflect general labour market dynamics and involve biases of the changing practices and habits regarding registering with the employment service. The interaction of the intervention setup and the unemployment register dynamics may also lead to wrong conclusions, as would be the case here with the share of long-term unemployed if entries to the unemployment register were not analysed here in parallel (see Fig. 5). In principle, the best single data source to use for answering the research questions on the labour market impacts would be micro-data on employment spells, which could then be aggregated in any way depending on the needs.

Declarations.

Acknowledgements
Not applicable.

Author contributions
The authors have contributed equally to the paper.

Funding
This paper was prepared as part of the project "Impacts of European Policies on Socio-Economic Development and Public Policies in Croatia (EUROIMPACT)" funded by the National Recovery and Resilience Plan 2021–2026.

Data availability
Two of the datasets generated and/or analysed during the current study are available in the ***Statistika online*** repository (https://statistika.hzz.hr/) and from the Ministry of Regional Development and EU funds (http://regionalni.weebly.com/uploads/5/8/0/0/58005979/vrijednosti_indeksa_razvijenosti_i_pokazatelja_za_izra%C4%8Dun_indeksa_razvijenosti_2017-2013-2010.xlsx).
The remaining data are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors have no financial or non-financial interests to disclose.

Received: 12 June 2025 / Accepted: 6 July 2026
Published online: 20 July 2026

References

Adriaenssens, S., Theys, T., Verhaest, D., Deschacht, N.: Subsidized household services and informal employment: the Belgian service voucher policy. J. Social Policy. **52**(1), 85–106 (2023)
Ben-Michael, E., Feller, A., Rothstein, J.: Synthetic Controls with Staggered Adoption. J. Royal Stat. Soc. Ser. B: Stat. Methodol. **84**, 351–381 (2022). https://doi.org/10.1111/rssb.12448
Brändle, T., Fervers, L.: Give it Another Try: What are the Effects of a Job Creation Scheme Especially Designed for Hard-to-Place Workers? J. Labor Res. **42**, 382–417 (2021). https://doi.org/10.1007/s12122-021-09322-x
Brown, A.J., Koettl, J.: Active labor market programs - employment gain or fiscal drain? IZA J. Labor Econ. **4**, 12 (2015). https://doi.org/10.1186/s40172-015-0025-5
Callaway, B., Sant'Anna, P.H.C.: Difference-in-Differences with multiple time periods. J. Econ. **225**, 200–230 (2021). (a) https://doi.org/10.1016/j.jeconom.2020.12.001
Callaway, B., Sant'Anna, P.H.C.: did: Difference in Differences. R package. https://cloud.r-project.org/web/packages/did/index.html, (2021)(b)
Callaway, B., Goodman-Bacon, A., Sant'Anna, P.H.C.: Difference-in-Differences with a Continuous Treatment, (2024). http://arxiv.org/abs/2107.02637
Card, D., Kluve, J., Weber, A.: What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. J. Eur. Econ. Assoc. **16**, 894–931 (2018). https://doi.org/10.1093/jeea/jvx028
Caswell D, Kupka P, Larsen F, Berkel R. van: The Frontline Delivery of Welfare-to-Work in Context. In: Berkel R. van Caswell D, Kupka P, and Larsen F. (eds.) Frontline Delivery of Welfare-to-Work Policies in Europe. pp. 1–13. Routledge (2017) https://doi.org/10.4324/9781315694474-1
R Core Team: R: A Language and Environment for Statistical Computing, (2024). https://www.r-project.org/
Council of the European Union: Council recommendation of 15 February 2016 on the integration of the long-term unemployed into the labour market: (2016)
Croatian Pension Insurance Fund: Statistics on the insured. https://www.mirovinsko.hr/hr/statistika/4319. Accessed 30 November 2025
Croatian Employment Service: Statistika online. (2025). https://statistika.hzz.hr/. Accessed 30
Croatian Bureau of Statistics: Population by Age, Sex and Type of Settlement, by Towns/Municipalities, 2021 Census. (2025). https://podaci.dzs.hr/media/td3jvrbu/popis_2021-stanovnistvo_po_gradovima_opcinama.xlsx. Accessed 30
Croatian Bureau of Statistics: Population estimate by sex, by towns/municipalities, 31 December. (2025). https://web.dzs.hr/PxWeb/pxweb/en/Stanovni%C5%A1tvo/Stanovni%C5%A1tvo__Procjene%20stanovni%C5%A1tva/SP31_2.px/. Accessed 30
Croatian Employment Service: Ipsos: Vanjska evaluacija mjera aktivne politike tržišta rada 2010.-2013.: sumarno evaluacijsko izvješće. Croatian Employment Service, Zagreb (2016)
Cseres-Gergely, Z., Molnár, G.: Labour market situation following exit from public works. In: Fazekas, K., Varga, J. (eds.) The Hungarian Labour Market 2015, pp. 148–159. Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budapest (2015)
Csillag, M.: European Support to vulnerable groups. Thematic paper for the European Network of Public Employment Services. European Commission (2021)

Dováľová, G., Hvozdíková, V.: The effects of public works programs - the case of the Slovak Republic. European Scientific Journal, ESJ. 10, (2014)
Escudero, V.: Are active labour market policies effective in activating and integrating low-skilled individuals? An international comparison. IZA J. Labor Policy. **7**, 4 (2018). https://doi.org/10.1186/s40173-018-0097-5
European Commision: An Agenda for new skills and jobs: A European contribution towards full employment: (2010)
European Commission: Public works: How can PES contribute toincreasing their value as an activation tool? Directorate-General for Employment. Social Affairs Inclusion, (2013). https://ec.europa.eu/social/BlobServlet?docId=13384&langId=en
European Commission, ICON Institute Public Sector GmbH: New forms of active labour market policy programmes. Publications Office, LU (2023)
European Training Foundation: Assessment of the effectiveness of active labour market policies in crisis and post-crisis situation. European Training Foundation (2022)
Eurostat: Labour Market Policy Database - Participants by LMP intervention - Croatia, (2025). https://webgate.ec.europa.eu/empl/redisstat/databrowser/view/lmp_partme$hr/default/table?lang=en&category=lmp_particip.lmp_particip_me
Fargion, V., Profeti, S.: The social dimension of Cohesion policy. In: Piattoni, S., Polverari, L. (eds.) Handbook on Cohesion Policy in the EU, pp. 475–490. Edward Elgar Publishing (2016) https://doi.org/10.4337/9781784715670
Ferreira, F.H.G., Robalino, D.A.: Social Protection in Latin America: Achievements and Limitations, (2010). https://papers.ssrn.com/abstract=1604349
Hujer, R., Thomsen, S.L.: How do the employment effects of job creation schemes differ with respect to the foregoing unemployment duration? Labour Econ. **171**, 38–51 (2010). https://doi.org/10.1016/j.labeco.2009.07.001
Jaenichen, U., Stephan, G.: The effectiveness of targeted wage subsidies for hard-to-place workers. Appl. Econ. **43**(10), 1209–1225 (2011). https://doi.org/10.1080/00036840802600426
Kluve, J.: The Effectiveness of European Active Labor Market Policy, (2006). https://papers.ssrn.com/abstract=892341
Kluve, J.: The effectiveness of European active labor market programs. Labour Econ. **17**, 904–918 (2010). https://doi.org/10.1016/j.labeco.2010.02.004
Kluve, J., Lehmann, H., Schmidt, C.M.: Active Labour Market Policies in Poland: Human Capital Enhancement, Stigmatization or Benefit Churning? (1999). https://papers.ssrn.com/abstract=155089
Koltai, L.: Work instead of social benefit? Public works in Hungary. In: GHK Consulting Ltd, CERGE-EI (eds.) Peer Review on Activation measures in times of crisis: the role of public works, pp. 1-13 . Riga, Latvia (2012) https://www.researchgate.net/publication/317402127_%27WORK_INSTEAD_OF_SOCIAL_BENEFIT_PUBLIC_WORKS_IN_HUNGARY%27
L'Horty, Y., Graceffa, S., Kerivel, A., Le Gallo, J., Lima, L., Petrella, F., Pin, C., Wolff, F.-C.: Deuxième évaluation de l'expérimentation Territoires zéro chômeur de longue durée - Note d'étape. France Stratégie and DARES (2024)
Larsen, C.A.: Policy paradigms and cross-national policy (mis)learning from the Danish employment miracle. J. Eur. Public. Policy. **9**, 715–735 (2002). https://doi.org/10.1080/13501760210162320
Layard, R., Nickell, S., Jackman, R.: Unemployment: Macroeconomic Performance and the Labour Market. Oxford University Press, Oxford, New York (2005)
Leduc, E., Tojerow, I.: Home work: Exploring the labor market effects of subsidizing domestic services. Labour Econ. **90** (2024). https://doi.org/10.1016/j.labeco.2024.102595
Malo, M.A.: Finding proactive features in labour market policies: a reflection based on the evidence, (2018)
Martin, J.P.: Activation and active labour market policies in OECD countries: stylised facts and evidence on their effectiveness. IZA J. Labor Policy. **4**, 4 (2015). https://doi.org/10.1186/s40173-015-0032-y
Martin, J.P., Grubb, D.: What Works and for Whom: A Review of OECD Countries' Experiences with Active Labour Market Policies. Swed. Econ. Policy Rev. **8**, 9–56 (2001). https://doi.org/10.2139/ssrn.348621
Matković, T., Babić, Z., Vuga, A.: Evaluacija mjera aktivne politike zapošljavanja 2009. i 2010. godine u Republici Hrvatskoj. Revija za socijalnu politiku. **19**, 303–336 (2012). https://doi.org/10.3935/rsp.v19i3.1100
Ministry of Regional Development and EU Funds: Vrijednosti indeksa razvijenosti i pokazatelja za izračun indeksa razvijenosti 2010., 2013. i 2017., (2018). http://regionalni.weebly.com/uploads/5/8/0/0/58005979/vrijednosti_indeksa_razvijenosti_i_pokazatelja_za_izra%C4%8Dun_indeksa_razvijenosti_2017-2013-2010.xlsx
Molnár, G., Bazsalya, B., Bódis, L., Kálmán, J.: Public works in Hungary: Actors, allocation mechanisms and labour market mobility effects. Socio hu Társadalomtudományi Sz. 117–142 (2019). https://doi.org/10.18030/socio.hu.2019en.116
Murgai, R., Ravallion, M., van de Walle, D.: Is Workfare Cost-effective against Poverty in a Poor Labor-Surplus Economy? World Bank Econ. Rev. **30**, 413–445 (2016)
Nestić, D., Tomić, I.: Jobless Population and Employment Flows in Recession. J. Balkan Near East. Stud. **20**, 273–292 (2018). https://doi.org/10.1080/19448953.2018.1385271
OECD: OECD Economic Surveys: Hungary 2014. OECD (2014)
OECD, Department of Social Protection, Ireland, European Commission, Joint Research Centre: Impact Evaluation of Ireland's Active Labour Market Policies. OECD: (2024)
Pompili, M., Kluve, J., Jessen, J., Seebauer, J., Gallassi, G., Peruccacci, E.: Meta-analysis of the ESF counterfactual impact evaluations: final report. Publications Office of the European Union (2023)
Porreca, Z.: Synthetic difference-in-differences estimation with staggered treatment timing. Econ. Lett. **220**, 110874 (2022). https://doi.org/10.1016/j.econlet.2022.110874
Raz-Yurovich L, Marx I. What does state-subsidized outsourcing of domestic work do for women's employment? The Belgian service voucher scheme. J. Eur. Soc. Policy. 28, 104–115 (2018). https://doi.org/10.1177/0958928717709173
Robalino, D.A., Newhouse, D., Rother, F.: Labor and Social Protection Policies during the Crisis and the Recovery. In: Banerji, A., Newhouse, D., Paci, P., Robalino, D. (eds.) Working through the Crisis: Jobs and Policies in Developing Countries during the Great Recession, pp. 85–134. The World Bank, Washington D.C. (2013)
Rodríguez-Planas, N., Jacob, B.: Evaluating active labor market programs in Romania. Empir. Econ. **38**, 65–84 (2010). https://doi.org/10.1007/s00181-009-0256-z
Sjögren, A.: How long and how much? Learning about the design of wage subsidies from policy changes and discontinuities. Labour Econ. **34**, 127–137 (2015). https://doi.org/10.1016/j.labeco.2015.03.009
Subbarao, K., del Ninno, C., Andrews, C., Rodríguez-Alas, C.: Public works as a safety net: design, evidence and implementation. The World Bank, Washington D.C. (2013)
Szabó, L.T.: The Effect of Public Work Programme in Hungary on Private Sector Wages, (2025). https://www.ssrn.com/abstract=5094028
Thomsen, S.L., Walter, T.: Temporary Extra Jobs for Immigrants: Merging Lane to Employment or Dead-End Road in Welfare? LABOUR. **24**, 114–140 (2010). https://doi.org/10.1111/j.1467-9914.2010.00505.x
Tomić, I., Zilic, I.: Working for 200 Euro? The Unintended Effects of Traineeship Reform on Youth Labor Market Outcomes. LABOUR. **34**, 347–371 (2020). https://doi.org/10.1111/labr.12176
Tzannatos, Z.: Women and Labor Market Changes in the Global Economy: Growth Helps, Inequalities Hurt and Public Policy Matters. World Dev. **27**, 551–569 (1999). https://doi.org/10.1016/S0305-750X(98)00156-9
Vooren, M., Haelermans, C., Groot, W., van den Maassen, H.: The Effectiveness of Active Labor Market Policies: A Meta-Analysis. J. Economic Surv. **33**, 125–149 (2019). https://doi.org/10.1111/joes.12269
Wolff, J., Stephan, G.: Subsidized work before and after the German Hartz reforms: design of major schemes, evaluation results and lessons learnt. IZA J. Labor Policy. **2**, 16 (2013). https://doi.org/10.1186/2193-9004-2-16
WYG: Vrednovanje Prioritetne osi 2 „Socijalno uključivanje: Završno izvješće o provedenom vrednovanju. Ministry of Labour, Pension System. Family and Social Policy & WYG, Zagreb (2021)

Publisher's note