Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Shortening unemployment benefit duration in jobless households with children: quasi-experimental evidence from Poland

Aleksandra Sojka

aleksandramariasojka@outlook.com

SGH Warsaw School of Economics

Research Article

Keywords: Unemployment insurance, Potential benefit duration, Benefit exhaustion, Jobless households, Difference-in-differences, Administrative data, Poland

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11641931/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Shortening unemployment benefit duration in jobless households with children: quasi-experimental evidence from Poland

Aleksandra Sojka[1]

SGH Warsaw School of Economics, Warsaw, Poland

Abstract

Whether shorter unemployment benefit entitlement speeds up the return to work is central to the design of unemployment insurance, yet little is known about the response of households in which no adult is employed. I study a Polish reform that, from February 2009, reduced the maximum potential benefit duration from 18 to 12 months for one statutory category of claimants: parents of a child below the age of 15 whose spouse was also registered as unemployed and had not yet lost benefit entitlement when the claimant's entitlement began. Using administrative records of 1.77 million benefit spells of parents drawn from the national unemployment register, I compare affected parents with other benefit-receiving parents in a difference-in-differences design, reinforced by reweighting, matching, individual fixed effects, doubly robust group-time estimation, formal sensitivity analysis and placebo tests. The reform reduced benefit collection by about 2.9 months – one fifth of the pre-reform mean – and raised the probability of exhausting entitlement by 10 percentage points. Recorded exits to employment did not change; confidence intervals exclude changes larger than about ± 3.5 percentage points. Estimated effects on registered unemployment duration fall from -3.6 months in repeated cross-sections to a statistically insignificant -1.3 months under within-individual identification, a gap attributable to the time-varying composition of the eligible inflow. For households facing multiple employment constraints, entitlement length operated as an insurance parameter rather than an activation lever.

[1]Corresponding author. SGH Warsaw School of Economics, al. Niepodległości 162, 02-554 Warsaw, Poland. E-mail: aleksandrasojka98@gmail.com. ORCID: 0000-0003-1667-5208.

Keywords: Unemployment insurance; Potential benefit duration; Benefit exhaustion; Jobless households; Difference-in-differences; Administrative data; Poland

JEL classification: C23; J13; J64; J65; J68

1 Introduction

The duration of unemployment benefit entitlement is one of the two central levers of unemployment insurance design. Together with the benefit level, it determines how much protection the system provides against earnings loss and how strongly it distorts the incentive to return to work (Baily, 1978; Chetty, 2006; Schmieder and von Wachter, 2016). A large quasi-experimental literature finds that longer potential benefit duration (PBD) extends unemployment spells and that exit rates from unemployment rise sharply as entitlement approaches exhaustion (Katz and Meyer, 1990; Meyer, 1990; Card et al., 2007a; Lalive, 2008). On this evidence, shortening entitlement is routinely advocated not only as a fiscal measure but as an activation instrument: a shorter horizon of income support should bring exits – and in particular exits to employment – forward.

Whether that reasoning extends to all groups of claimants is far less clear. The incentive mechanism presumes that the binding constraint on re-employment is the margin that benefit duration targets: search effort and the reservation wage. For households in which both adults are out of work and young children are present, that presumption is questionable in both directions. On the one hand, such households have no labour earnings, so the insurance value of benefits is plausibly high, and liquidity effects could make measured duration responses large (Chetty, 2008; Card et al., 2007b). On the other hand, caregiving obligations and the absence of a second earner constrain which jobs are feasible at all, so a shorter benefit horizon may simply end transfers earlier without moving the employment margin. Existing heterogeneity analyses are organised mainly around age, gender and labour-market conditions (Lalive, 2007, 2008; Le Barbanchon, 2016; Schmieder and von Wachter, 2016; Lopes, 2022), and the literature that connects unemployment insurance to parenthood studies benefit levels and birth-related outcomes rather than duration responses (Kessler and Hevenstone, 2022). Direct evidence on how jobless households with children respond to entitlement length is scarce, although these are precisely the households for which many systems provide targeted extensions and for which the consequences of unemployment extend to children (Rege et al., 2011; Feather, 1990).

This paper provides such evidence from a sharp institutional reform. Under Article 73 of the Polish Act of 20 April 2004 on the Promotion of Employment and Labour Market Institutions,

unemployed parents caring for at least one child below the age of 15 were entitled to 18 months of unemployment benefits – rather than the standard 6 or 12 – if their spouse was also registered as unemployed and had not yet lost his or her own benefit entitlement on the date the claimant's entitlement began (Employment Promotion and Labour Market Institutions Act, 2004). The amendment of 19 December 2008, in force from 1 February 2009, cut this group's maximum entitlement from 18 to 12 months (Act of 19 December 2008, 2009). No other parameter of the group's benefits changed: eligibility conditions, the benefit-level schedule and contribution requirements remained as before. The reform thus shifted a single design margin – potential benefit duration – for a well-defined statutory category of claimants drawn from jobless two-adult households with children.

I estimate the causal effects of this entitlement cut on five outcomes: the duration of benefit collection, the duration of registered unemployment, benefit exhaustion, the fraction of entitlement used, and recorded exits from the register to employment. The analysis uses individual-level administrative records from the national unemployment register covering 1.77 million benefit spells of parents whose entitlement began between June 2004 and December 2019, of which 2,723 were granted under the parental extension. Identification rests on a difference-in-differences (DiD) comparison between the affected parents and other benefit-receiving parents of children below 15, who experienced the same calendar shocks and the same concurrent reform of the regional entitlement criterion but whose own person-based entitlement rules did not change. The baseline two-way fixed-effects estimates are reinforced by propensity-score reweighting and matching, individual fixed effects identified from repeat spells, group-specific trends, the doubly robust group-time estimator of Callaway and Sant'Anna (2021), formal sensitivity analysis with respect to violations of parallel trends (Rambachan and Roth, 2023), wild cluster bootstrap inference (Cameron et al., 2008), and a comparison of adjacent claim cohorts with a placebo counterpart in the spirit of Johnston and Mas (2018).

Three findings emerge, and distinguishing how firmly each is established is part of the contribution. First, the response of benefit receipt is large, precisely estimated and robust across every estimator: benefit collection fell by 2.9–3.1 months – about one fifth of the pre-reform treated mean of 14.7 months, or roughly half a month per month of entitlement removed – and

the probability of exhausting the entitlement rose by about 10 percentage points. Part of this response is mechanical, since spells longer than the new cap were truncated by statute; its size nevertheless shows that the abolished six months of entitlement had been heavily used. Second, recorded exits to employment did not change. The estimate is close to zero in every specification, and the baseline confidence interval excludes changes larger than about ± 3.5 percentage points against a pre-reform mean of 60%. Third, the effect on registered unemployment duration is estimator-dependent in an informative way: repeated cross-sections indicate a decline of about 3.6 months, but the estimate falls to a statistically insignificant -1.3 months when identification uses only within-individual variation. Because eligibility for the parental extension depends on the spouse's labour-market state, the composition of the eligible inflow co-moves with labour-market conditions, and estimators that pool repeated cross-sections attribute part of that compositional change to the reform.

The paper makes three contributions. First, it adds direct evidence on a margin of heterogeneity that the unemployment insurance literature has rarely been able to isolate: the response of households with no employed adult and dependent children to the length of entitlement. The estimated pattern – a strong compression of benefit receipt combined with no detectable employment response – contrasts with the average pattern in the international evidence, in which duration cuts raise job-finding rates (Filges et al., 2018; Schmieder and von Wachter, 2016), and with Polish discontinuity-based estimates for the general claimant population, in which PBD changes move unemployment duration itself (Jessen et al., 2026). Second, the analysis carries a methodological lesson for the evaluation of group-targeted entitlements. Where eligibility depends on household labour-market states, the eligible population changes with the business cycle; the wedge documented here between cross-sectional and within-individual estimates shows that designs which ignore this composition margin can misattribute compositional change to behaviour. Third, the results speak directly to current policy. Poland's 2025 Act on the Labour Market and Employment Services retained a 12-month entitlement for claimants in analogous family circumstances while abolishing the geographic differentiation of benefit duration (Labour Market and Employment Services Act, 2025); the present estimates imply that the length of such group-targeted extensions governs primarily the duration of income support,

not the speed of labour-market reintegration.

The remainder of the paper is organised as follows. Section 2 develops the conceptual framework and reviews the related literature. Section 3 describes the institutional setting and the reform. Section 4 presents the data, sample and descriptive evidence. Section 5 sets out the empirical strategy and its identifying assumptions. Section 6 reports the main results, Section 7 the robustness and heterogeneity analysis, and Section 8 discusses interpretation and limitations. Section 9 concludes.

2 Conceptual framework and related literature

2.1 The duration margin in the theory of unemployment insurance

The modern analysis of unemployment insurance starts from the trade-off between consumption smoothing and search incentives. Because private insurance against unemployment fails under adverse selection and moral hazard, publicly provided benefits raise welfare by stabilising consumption, but they also raise the value of remaining unemployed (Baily, 1978; Atkinson and Micklewright, 1991; Krueger and Meyer, 2002). In the canonical formulation, the optimal generosity of benefits equates the marginal consumption-smoothing gain to the marginal incentive cost, summarised by the consumption drop at unemployment and the elasticity of unemployment duration with respect to generosity (Baily, 1978; Chetty, 2006). An important refinement distinguishes why benefits prolong unemployment: Chetty (2008) shows that a substantial part of measured duration responses reflects a liquidity effect – cash-on-hand allowing constrained households to search less desperately – rather than a substitution effect of distorted incentives, with the two having very different welfare implications (see also Card et al., 2007b).

Job-search theory links the duration margin to behaviour over the spell. With finite entitlement, the value of unemployment declines as exhaustion approaches, so reservation wages fall and search intensity rises, generating an increasing re-employment hazard towards the exhaustion point (Mortensen, 1977; Nickell, 1979). A reduction in PBD moves this incentive profile forward: exits should rise around the new, earlier exhaustion point. Two further results qualify the welfare reading of such responses. More generous insurance can improve the allocation

of workers to jobs by subsidising selective search (Acemoglu and Shimer, 1999; Marimon and Zilibotti, 1999), although prolonged unemployment may also erode skills (Ljungqvist and Sargent, 1998); and the timing of generosity matters, with evidence that the moral-hazard cost of benefits is, if anything, larger when benefits are paid early in the spell (Kolsrud et al., 2018). Surveys of this literature identify the benefit level and potential duration as the two central design margins and document that duration responses are concentrated among claimants with long spells (Schmieder and von Wachter, 2016; Tatsiramos and van Ours, 2014; Le Barbanchon et al., 2024).

For the population studied here, the theory makes no unambiguous prediction. Jobless two-adult households with children have no labour earnings, so their liquidity position suggests strong sensitivity to transfer income; at the same time, if caregiving and the absence of feasible job offers – rather than reservation wages – bind, the incentive channel that a shorter PBD is meant to activate has little to operate on. Which margin dominates is an empirical question, and the answer determines whether a duration cut works as an activation instrument or only as a reduction in insurance.

2.2 Quasi-experimental evidence on potential benefit duration

The empirical literature consistently finds that benefit generosity extends measured unemployment durations, with pronounced behavioural responses near exhaustion. Classic evidence from the United States shows higher exit rates as benefits run out (Meyer, 1990; Katz and Meyer, 1990); later work using sharp institutional variation in Austria, Norway and elsewhere corroborates both the duration response and the exhaustion spike (Lalive et al., 2006; Røed and Zhang, 2003; Card et al., 2007a). The interpretation of the spike requires care: part of it reflects exits from the unemployment register that do not correspond to job starts, so that outcomes defined on benefit receipt, registered unemployment and total non-employment can respond differently to the same reform (Card et al., 2007a; Schmieder and von Wachter, 2016). This measurement distinction is central to the present paper, which estimates effects on benefit collection and registered unemployment separately and treats recorded exit destinations as administrative declarations rather than verified employment states.

Evidence on duration cuts specifically supports an average incentive effect. Pooling quasi-experimental evaluations, Filges et al. (2018) find that reductions in maximum entitlement raise the hazard of exit to employment by roughly 10% on average, with little evidence of systematic deterioration in re-employment wages. Studies of individual reforms find substantial responses of benefit receipt and unemployment duration to entitlement length (Lalive, 2008; Le Barbanchon, 2016; Johnston and Mas, 2018), and the survey by Lopes (2022) places the elasticity of unemployment duration with respect to PBD in a moderate positive range. Two caveats temper these averages. First, meta-analytic work correcting for publication selection implies that headline disemployment effects in the published record overstate the true average, by roughly a factor of two in preferred estimates (Cohen and Ganong, 2024). Second, effects on match quality are modest in either direction: exploiting an age discontinuity in Austria, Nekoei and Weber (2017) find that longer entitlement slightly improves re-employment wages through better employer matches, an effect partly offset by the wage cost of longer non-employment. Welfare analysis of duration reforms additionally depends on benefit levels and schedules (Landais, 2015; Kolsrud et al., 2018).

Against this background, the contribution of the present setting is the population to which the variation applies. Heterogeneity in duration responses has been documented by age, gender and local labour-market conditions (Lalive, 2007, 2008; Schmieder and von Wachter, 2016), but statutory variation that isolates households with two non-employed adults and dependent children is rare. Such households are of policy interest in their own right: many systems grant them extended entitlements, their exposure to child-welfare consequences of unemployment is direct (Rege et al., 2011), and – as the results below show – their behavioural response need not resemble the population average.

2.3 The Polish evidence

Poland offers unusually clean rule-based variation in benefit duration. The system that emerged from the post-1989 transition (Boeri, 2000) pairs strict eligibility with low, flat-rate benefits: net replacement rates are among the lowest in the OECD (OECD, 2024), and fewer than one in five registered unemployed receive benefits at any point in time (Kalinowska, 2016). Early

transition-era studies found muted average effects of entitlement reforms – Puhani (2000) detects no significant change in unemployment duration after the 1991 move to time-limited benefits – while microeconometric work documented higher exit rates as exhaustion approaches (Adamchik, 1999), foreshadowing the incentive patterns found elsewhere.

Recent work exploits the distinctive Polish rule that ties standard PBD to the county unemployment rate, a design made meaningful by the persistence of local labour-market disparities (Tyrowicz and Wójcik, 2010). Using the universe of registered spells and the cross-county discontinuity, Jessen et al. (2026) find that extending PBD from 6 to 12 months raises the unemployment duration of directly affected claimants below age 50 by about 13%, and separate individual responses from county-level spillovers. Jessen et al. (2025) show that benefit levels and durations interact, with longer entitlement amplifying the distortionary effect of higher benefit levels, including on inflows from employment. Quasi-experimental evidence on other Polish social-insurance margins points in the same direction of modest but detectable behavioural responses (Komada et al., 2019). Two features distinguish the present study within this literature: the treated and control populations are drawn from the same demographic group (benefit-receiving parents of young children), and the data permit within-individual identification from repeat spells, so that the sensitivity of cross-sectional estimates to composition can be assessed directly. As in this literature generally, the estimates concern claimants; a margin not studied here is take-up, which recent evidence shows to be far from universal among the eligible (Lachowska et al., 2025; McQuillan and Moore, 2025).

3 Institutional background

3.1 The Polish unemployment benefit system

Unemployment benefits were introduced in Poland at the outset of the transition (Employment Act, 1989) and acquired their lasting shape in the 1990s: the Act of 16 October 1991 replaced the initial earnings-related formula with a flat-rate benefit and a statutory time limit (Employment and Unemployment Act, 1991), and the Act of 14 December 1994 made the duration of entitlement depend on the county unemployment rate relative to the national average (Employ-

ment and Counteracting Unemployment Act, 1994). The Act of 20 April 2004 on the Promotion of Employment and Labour Market Institutions consolidated this architecture and governed the system throughout the period studied here (Employment Promotion and Labour Market Institutions Act, 2004).

Under the 2004 Act, entitlement required at least 365 days of contributory employment, with earnings at least equal to the minimum wage, within the 18 months preceding registration (Employment Promotion and Labour Market Institutions Act, 2004, art. 71). The benefit is essentially flat-rate: a nationally uniform base amount, indexed annually, paid at 80%, 100% or 120% depending on the claimant's total qualifying period. It is financed by the Labour Fund, a national fund fed chiefly by employer payroll contributions. Receipt is conditional on registered unemployment at a county labour office, availability for work and periodic reporting; refusal of a suitable job or programme offer without good cause ends eligibility. By international standards the system combines strict access with low generosity (Kalinowska, 2016; OECD, 2024).

3.2 Potential benefit duration and the parental extension

Potential benefit duration is the margin on which Polish law has differentiated most explicitly, and Table 1 summarises the statutory rules across successive regimes. In the original wording of Article 73 of the 2004 Act, the standard duration was 6 months in counties whose unemployment rate on 30 June of the preceding year did not exceed 125% of the national average, and 12 months in counties above this threshold. Independently of location, claimants aged 50 or above with at least 20 years of qualifying period received 12 months. The longest tier, 18 months, applied to two groups: claimants in counties exceeding 200% of the national average who also had 20 years of qualifying period, and – the rule studied here – claimants supporting at least one child below the age of 15 whose spouse was also registered as unemployed and had not yet lost his or her own benefit entitlement on the date on which the claimant acquired the right (Employment Promotion and Labour Market Institutions Act, 2004, art. 73). The parental extension thus selected households in which both spouses were simultaneously unemployed, the second spouse recently enough to retain entitlement – in effect, households with no labour

earnings at the moment the claimant's benefit began.

Table 1: Potential unemployment benefit duration (PBD) under successive Polish legal regimes, 2004–2025

Claimant's situation	2004–09	2009–14	2014–25	2025–
General rule (county at or below regional threshold)[a]	6 mo	6 mo	6 mo	6 mo
County above regional unemployment threshold[a]	12 mo	12 mo	12 mo	–
County >200% of national average and ≥20 years qualifying period	18 mo	–	–	–
Aged 50+ with ≥20 years qualifying period	12 mo	12 mo	12 mo	12 mo
Parent of child under 15 with unemployed spouse who had not lost benefit entitlement when the claimant's entitlement began[b]	**18 mo**	**12 mo**	**12 mo**	**12 mo**
Single parent[b]	–	–	12 mo	12 mo
Person with disability	–	–	–	12 mo
Member of large family (3+ children)	–	–	–	12 mo

Notes: 6 mo = 180 days, 12 mo = 365 days, 18 mo = 540 days. Column ranges denote the periods in which each regime applied: the original 2004 Act (in force from 1 June 2004), its amendments in force from 1 February 2009 and 27 May 2014, and the 2025 Act (in force from 1 June 2025). Legal bases: Art. 73 of the 2004 Act and its amendments (Employment Promotion and Labour Market Institutions Act, 2004; Act of 19 December 2008, 2009; Act of 14 March 2014, 2014); Art. 225 of the 2025 Act (Labour Market and Employment Services Act, 2025). The rule studied in this paper is set in bold. [a] Regional criterion: 12 months where the county unemployment rate on 30 June of the preceding year exceeded 125% (2004–2009) or 150% (2009–2025) of the national average; abolished by the 2025 Act. The time-degressive benefit amount in force from 1 January 2010 (a higher rate for the first 90 days) concerns the benefit level, not its duration, and is not shown. [b] Qualifying child aged under 15 up to 2025; from the 2025 Act, under 18 (24 if the child has a disability and continues education). The single-parent category was introduced by the 2014 amendment. A maternity extension applies throughout.

3.3 The 2009 reform

The amendment of 19 December 2008, in force from 1 February 2009, restructured Article 73 in two respects (Act of 19 December 2008, 2009). First, it eliminated the 18-month tier altogether: the parental extension was reduced from 18 to 12 months, uniformly across counties. Second, it raised the regional threshold from 125% to 150% of the national average unemployment rate, which reclassified a substantial number of counties from the 12-month to the 6-month standard category. Transitional provisions preserved acquired rights: benefits granted before 1 February 2009 ran their course under the old rules, while benefits granted from that date onward were subject to the new entitlement structure. The benefit-level schedule (80%, 100%

or 120% of the base amount) was not altered. A time-degressive profile of the base amount – a higher rate for the first 90 days of the spell – was introduced separately from January 2010 and applied to treated and control claimants alike. Later legislation is relevant mainly for external validity: the 2014 amendment added single parents to the 12-month category (Act of 14 March 2014, 2014), and the 2025 Act on the Labour Market and Employment Services abolished the geographic criterion entirely while retaining 12-month extensions for claimants in analogous family circumstances (Labour Market and Employment Services Act, 2025).

Two features of this setting matter for the empirical design. First, the parental extension is the only entitlement rule that changed exclusively for the treated group; all other changes introduced by the amendment (the regional threshold) or by contemporaneous legislation (the degressive profile) applied to treated and control parents symmetrically, conditional on county. Second, the reform was adopted six weeks before it entered into force, creating a window during which prospective claimants could still register under the old rules. Behavioural responses within this announcement window are documented in Appendix E and addressed in the robustness analysis.

4 Data and descriptive evidence

4.1 Data source and analysis sample

I use individual-level administrative records from the national unemployment register maintained by the public employment services. The extract underlying this study is restricted to registered unemployed persons who can be identified as parents: first, individuals whose benefit was granted under the parental extension and, second, other registered unemployed individuals for whom at least one child could be identified because household members were reported at the labour office, typically for insurance coverage. The unit of observation is the registered-unemployment spell. The extract comprises 7,409,695 spells of 2,419,816 individuals; it serves to construct spell histories and identify children, while the estimation sample is the much narrower benefit-recipient subset defined below. For each spell the register records exact start and end dates, the start and end dates and statutory basis of benefit receipt, and the reason for re-

moval from the register, together with demographic characteristics (date of birth, sex, education, place of residence and its urban/rural status), total prior work experience, occupation, and the birth dates of household members, from which the number and ages of children at registration are derived. County identifiers link each spell to the local unemployment rate published by Statistics Poland and hence to the statutory county-level PBD category in force in the year the benefit was granted.

Treatment status is observed directly rather than imputed: benefits granted under the parental extension carry a separate statutory basis code in the register. The treated group therefore consists of spells in which the benefit was granted under the extended-duration provision for parents of a child below 15 with an unemployed spouse who had not yet lost his or her own benefit entitlement when the claimant's began.[2] The control group consists of all other benefit-receiving parents with at least one child below the age of 15 at registration.

One selection property of this construction should be acknowledged openly. Control parents are identifiable as parents only because a child was reported at the labour office, so the control group covers benefit-receiving parents who disclosed a child rather than all benefit-receiving parents. For the difference-in-differences design this selection is not, by itself, a threat: identification requires only that the outcomes of the two groups would have moved in parallel in the absence of the reform, not that the control group be representative of all parents. Selection would bias the estimates only if the propensity to report a child changed differentially around the reform date, for which there is no institutional reason.

The analysis sample is restricted to spells with benefit entitlement starting between 1 June 2004 and 31 December 2019. The lower bound corresponds to the entry into force of the 2004 Act, which first codified the parental extension in the form described above; the upper bound ensures that outcomes are observed over an adequate follow-up horizon and are unaffected by the administrative disruption of the COVID-19 period. The resulting sample comprises 1,768,767 benefit spells, of which 2,723 are treated (1,598 with benefit start before February 2009 and 1,125 after). A small number of spells with incomplete covariate information drop

[2] A small number of records (1,061 in the analysis window) carry a continuation code associated with the same statutory basis, recorded when an existing entitlement was administratively re-registered rather than newly granted. Because the timing of the underlying grant cannot be established, these records are excluded from the analysis sample altogether.

out of estimation, so the regression samples comprise 1,767,335 spells, 2,722 of them treated. Roughly two thirds of sampled individuals appear with a single benefit spell (66.7%), while the remainder contribute repeated spells; the individual fixed-effects specifications are identified from 1,009,166 spells of 381,672 individuals observed at least twice, including 1,021 individuals who ever hold a treated spell and together contribute 2,712 spells.

The small size of the treated group relative to the register follows directly from the eligibility conditions, which require the simultaneous unemployment of both spouses with the spouse's entitlement still active. The annual inflow into treated status also varies strongly over time – from over 900 new treated claims in 2004 to 74 in 2008, recovering to 100–160 per year after 2009 – plausibly because the double-unemployment condition is satisfied more often when labour-market conditions are weak. The implications of this time-varying composition for identification are taken up in Section 5.2.

4.2 Outcome variables

Five outcomes are examined; all durations are expressed in months. *Benefit collection duration* is the length of benefit receipt recorded in the register; because the reform cut the maximum entitlement, an effect on this outcome combines the mechanical truncation of long spells with any behavioural change in how long claimants remain on benefits. *Unemployment duration* is the length of the registered unemployment spell from registration to removal from the register; it has no statutory cap, so an effect on it cannot be mechanical. Spells still open at the end of the observation window (1.2%) are retained with a censoring indicator. *Benefit exhaustion* is an indicator equal to one if recorded benefit receipt reaches the statutory entitlement (up to a few days of administrative tolerance), where the statutory entitlement is 18 months pre-reform and 12 months post-reform for treated spells.[3] For control spells the statutory entitlement follows the county category (6 or 12 months) and the age-50/20-year rule, the only person-based extension applicable to the control group. *Fraction of entitlement used* is the ratio of recorded benefit receipt to the statutory entitlement. *Exit to employment* is an indicator equal to one if the recorded reason for removal from the register is job take-up or business start-up.

[3] Statutory durations are defined in 30-day months, so the statutory "18 months" correspond to slightly less than 18 calendar months; this is why the observed mass points of benefit receipt lie just below the nominal caps.

One property of the destination measures should be stated at the outset: exit reasons are recorded at the moment of deregistration and reflect the reason declared to or established by the employment office at that moment. The register contains no information on activity after removal. All destination-based outcomes, including exit to employment, are therefore administrative records of the stated reason for leaving the register rather than verified post-exit labour-market states, and they carry no information on whether, for how long, or at what wage the reported employment was actually held. A disaggregation of all exit destinations is reported in Section 6.3.

4.3 Descriptive statistics

Table 2 reports means of covariates and outcomes by treatment group and reform period; Appendix C adds standard deviations and medians. Treated and control parents are similar in age (approximately 35 years at registration) and, before the reform, in work experience (10.9 versus 10.3 years). They differ in several respects: treated parents are less likely to reside in urban areas (45% versus 55% before the reform), have more children (1.83 versus 1.70) and younger children (mean age of the youngest child 5.4 versus 6.1 years), and are more concentrated in basic vocational education (46% versus 39%) with a lower share of tertiary education (4% versus 8%). Both groups exhibit a marked upward shift in educational attainment after 2009, consistent with secular trends in the registered unemployed population. The share of spells located in counties with a 12-month standard PBD falls in both groups after the reform – from 37% to 4% among the treated and from 51% to 38% among the controls – reflecting the simultaneous increase in the regional threshold; the sharper decline among the treated reflects their concentration in counties close to the previous threshold. These differences motivate the covariate adjustment and the propensity-score-based estimators introduced below.

The raw cell means already display the expected ordering. Before the reform, treated parents collected benefits for 14.7 months on average, close to the 18-month cap, against 5.9 months in the control group; their registered unemployment spells lasted 24.1 months against 11.7. After the reform, mean benefit collection among the treated fell to 10.4 months and mean unemployment duration to 19.4 months, while the corresponding control means changed only modestly

Table 2: Sample characteristics by treatment status and reform period

	Control parents		Treated parents	
	Before Feb 2009	After Feb 2009	Before Feb 2009	After Feb 2009
Covariates				
Female	0.559	0.612	0.501	0.602
Age at registration	34.67	35.44	34.88	34.79
Work experience (years)	10.25	9.35	10.90	9.18
Education: primary or less	0.208	0.161	0.216	0.160
Education: basic vocational	0.392	0.297	0.456	0.298
Education: secondary general	0.075	0.104	0.063	0.107
Education: secondary vocational	0.197	0.199	0.195	0.224
Education: post-secondary	0.050	0.066	0.030	0.077
Education: tertiary	0.078	0.172	0.041	0.134
Urban residence	0.546	0.531	0.451	0.517
Number of children	1.70	1.79	1.83	1.87
Age of youngest child	6.14	5.87	5.36	4.86
County PBD = 12 months	0.514	0.377	0.372	0.043
Outcomes				
Benefit collection duration (months)	5.92	5.24	14.74	10.42
(SD)	(3.84)	(3.50)	(4.12)	(2.26)
Unemployment duration (months)	11.73	11.22	24.06	19.40
(SD)	(14.24)	(13.42)	(18.66)	(14.43)
Benefit exhausted	0.411	0.390	0.485	0.568
Fraction of entitlement used	0.668	0.658	0.831	0.869
(SD)	(0.366)	(0.390)	(0.232)	(0.188)
Exit to employment	0.671	0.667	0.601	0.592
Censored spell	0.001	0.017	0.000	0.026
Spells	580,631	1,185,413	1,598	1,125

Notes: Analysis sample: benefit spells of parents of at least one child below age 15, benefit start June 2004–December 2019. Treated: benefits granted under the extended-duration provision for parents with an unemployed spouse. "Before/After Feb 2009" refers to the benefit start date relative to the entry into force of the 2009 amendment. Standard deviations reported for duration outcomes. A detailed version of this table, with standard deviations and medians for all continuous variables, is reported in Appendix C. Source: administrative records of Polish county labour offices; author's elaboration.

(5.2 and 11.2 months). The share of treated spells exhausting the entitlement rose from 49% to 57% despite the shorter entitlement, and the mean fraction of entitlement used rose from 0.83 to 0.87; in the control group both measures were stable. The share of spells ending in employment was approximately 0.60 in the treated group both before and after the reform.

Figure 1 plots Kaplan–Meier survival functions of benefit collection for treated and control spells, before and after the reform. The statutory caps are visible directly: the pre-reform treated

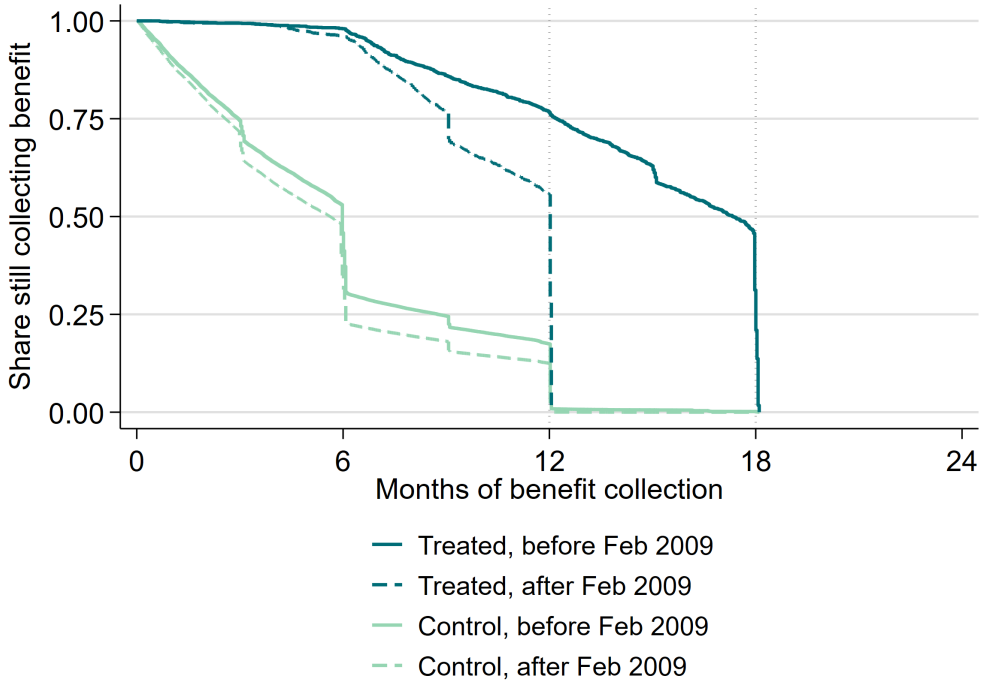

Figure 1: Kaplan–Meier survival functions of benefit collection, by treatment group and reform period

Notes: Survival functions estimated by the method of Kaplan and Meier (1958). Vertical lines at 12 and 18 months. Source: administrative records of Polish county labour offices; author's elaboration.

curve collapses at 18 months, the post-reform treated curve at 12, while the control curves lie far below throughout. The location of the upper mass point shifts exactly as implied by the entitlement change. Among pre-reform treated spells with recorded durations between 17.4 and 18.1 months, 92% terminate within a few days of the statutory maximum, and among post-reform spells the corresponding concentration at twelve months is 96%. The pronounced mass point at exhaustion in both regimes is consistent with the exhaustion-spike literature (Katz and Meyer, 1990; Card et al., 2007a).

Figure 2 shows the corresponding survival functions of registered unemployment. The post-reform treated survival function lies below the pre-reform function from approximately the twelfth month of the spell onward, with the steepest divergence in the 12–18 month interval – between the new and the old exhaustion point – while the control curves show only a modest pre–post difference. The share of completed treated spells ending between the twelfth and eighteenth month rises from 22.4% before the reform to 33.3% after it, whereas in the control group the corresponding share declines (17.1% to 14.0%). The descriptive evidence thus locates

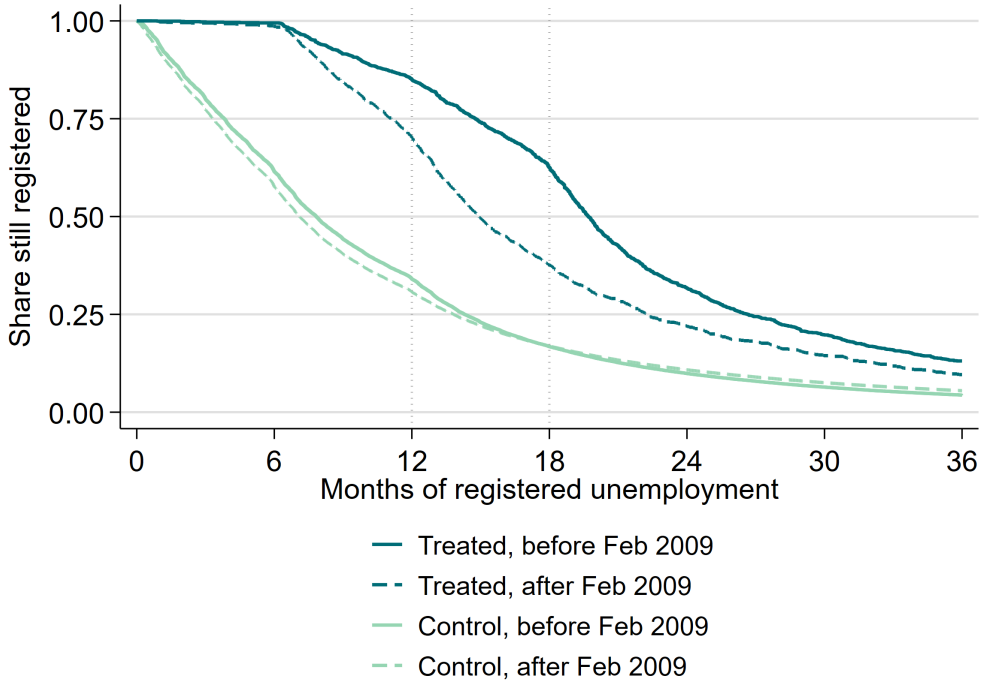

Figure 2: Kaplan–Meier survival functions of registered unemployment, by treatment group and reform period

Notes: Survival functions estimated by the method of Kaplan and Meier (1958). Vertical lines at 12 and 18 months. Source: administrative records of Polish county labour offices; author's elaboration.

the change in exit behaviour precisely in the region of the outcome distribution affected by the entitlement change.

4.4 Raw group trends

Before turning to the formal design, the raw data can show whether the two groups moved in parallel in the years preceding the reform. Figures 3 and 4 plot unadjusted yearly means of the five outcomes by treatment group. Three features stand out. First, the level gap between the groups is large and persistent, reflecting the longer entitlement and the particular circumstances of the treated group: treated parents collect benefits roughly nine months longer and remain registered roughly twelve months longer. This permanent gap is what the group term in the difference-in-differences design absorbs. Second, before 2009 the two series move broadly in parallel for benefit collection and the entitlement-based outcomes – the visual counterpart of the parallel-trends assumption assessed formally in Section 6.2. Third, at the reform date the treated series shift sharply – benefit collection falls towards the new twelve-month cap and the

exhaustion share rises – while the control series continue along their gradual path. The task of the empirical design is to turn this visual contrast into an estimate that nets out composition and common shocks.

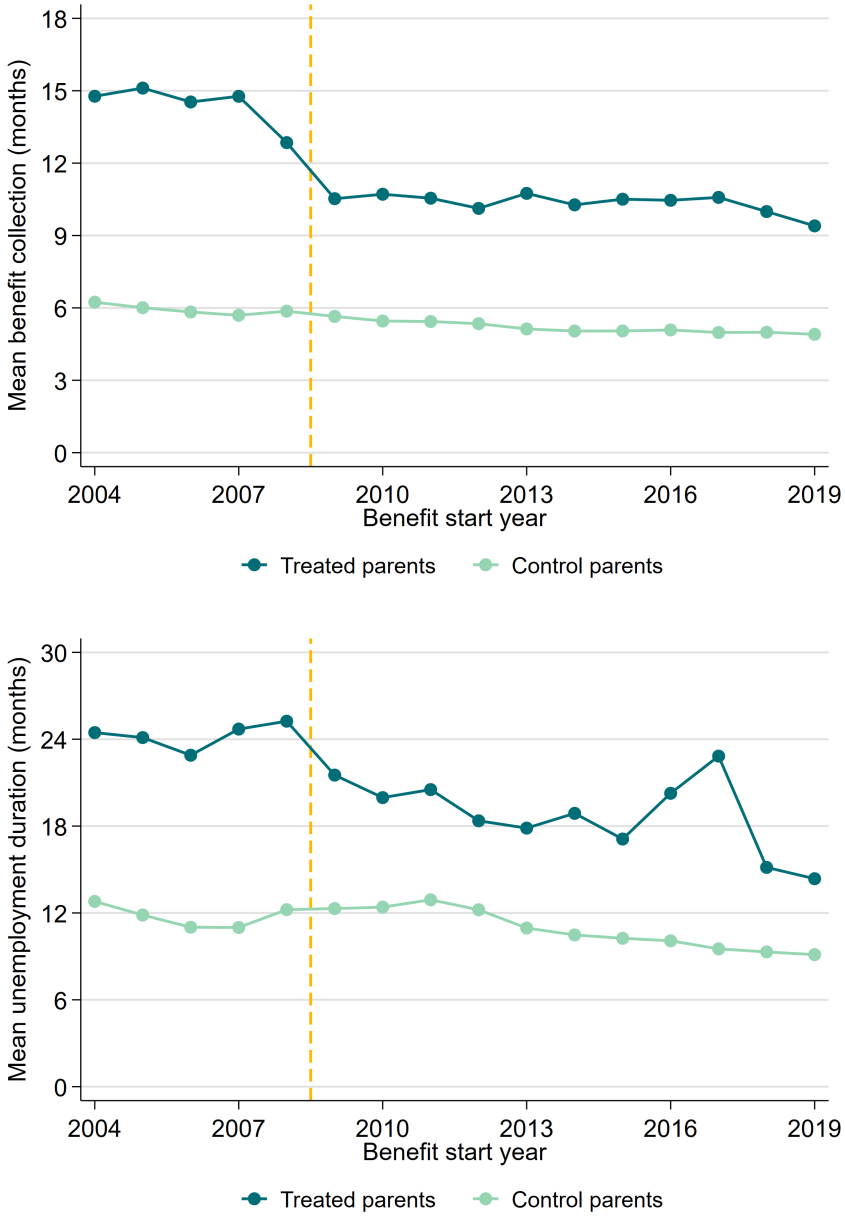

Figure 3: Raw yearly means by treatment group: duration outcomes

Notes: Benefit collection (top) and registered unemployment duration (bottom), unadjusted yearly means by treatment group and benefit-start year. The dashed vertical line marks the 2009 reform. Source: administrative records of Polish county labour offices; author's elaboration.

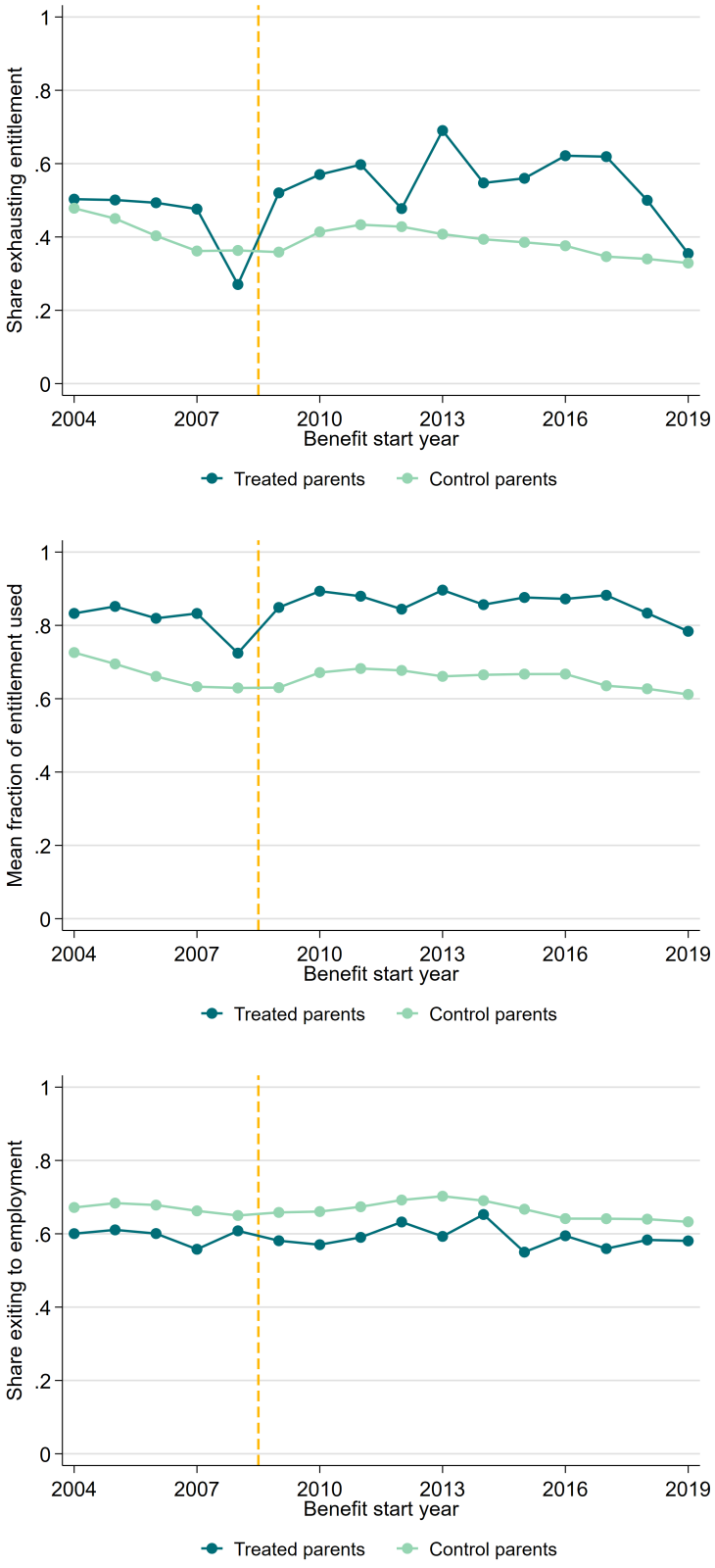

Figure 4: Raw yearly means by treatment group: benefit exhaustion, entitlement use and employment exits

Notes: Benefit exhaustion (top), fraction of entitlement used (middle) and exit to employment (bottom), unadjusted yearly means by treatment group and benefit-start year. The dashed vertical line marks the 2009 reform. Source: administrative records of Polish county labour offices; author's elaboration.

5 Empirical strategy

5.1 Baseline difference-in-differences

Difference-in-differences identifies a treatment effect by contrasting the change in outcomes of a treated group over the treatment date with the contemporaneous change in a control group.[4] In the canonical two-group, two-period case the estimand is

$$\beta_{\text{DiD}} = \left(\bar{Y}^{T}_{\text{post}} - \bar{Y}^{T}_{\text{pre}}\right) - \left(\bar{Y}^{C}_{\text{post}} - \bar{Y}^{C}_{\text{pre}}\right), \tag{1}$$

where $\bar{Y}^{g}_{p}$ denotes the mean outcome of group $g \in \{T, C\}$ in period p. The first difference removes all time-invariant differences between the groups; the second removes all changes over time common to both. In the present application the treated group consists of parents entitled to the extended benefit duration, the control group of other benefit-receiving parents of children below 15, and the treatment date is the entry into force of the amendment on 1 February 2009, which reduced the treated group's entitlement from 18 to 12 months while leaving the control group's entitlement rules unchanged.

In its simplest regression form, the estimand of (1) is recovered from

$$Y_{it} = \alpha + \gamma\, T_i + \delta\, Post_t + \beta\,(T_i \times Post_t) + \varepsilon_{it}. \tag{2}$$

I do not estimate equation (2) in this form. With sixteen years of monthly benefit inflows, a single $Post_t$ dummy would compress the entire time path of Polish labour-market conditions into one step change and leave seasonal and other time-varying conditions in the error term. The baseline specification therefore replaces the single post indicator with a full set of benefit-start month-by-year fixed effects and adds covariates and county fixed effects:

$$Y_{ict} = \beta\,(T_i \times Post_t) + \gamma\, T_i + X'_{it}\delta + \mu_c + \lambda_t + \varepsilon_{ict}, \tag{3}$$

[4] The design has a long lineage in labour economics, going back to the training-programme evaluations of Ashenfelter and Card (1985); textbook treatments are provided by Angrist and Pischke (2009), and recent methodological surveys by Roth et al. (2023) and Goodman-Bacon (2021).

where Y_{ict} is an outcome of spell i in county c with benefit start in calendar month t; T_i equals one for benefits granted under the parental extension; $Post_t$ equals one for benefit start on or after 1 February 2009; X_{it} contains sex, age at registration, work experience, education (six categories), urban residence, number of children, and age of the youngest child; μ_c are county fixed effects; and λ_t are benefit-start month-by-year fixed effects, which absorb the main effect of $Post_t$, common seasonality, and the aggregate path of labour-market conditions. Standard errors are clustered at the county level (329 clusters); inference is additionally verified with the wild cluster bootstrap (Cameron et al., 2008; Roodman et al., 2019).

The parameter of interest β identifies the average treatment effect on the treated (ATT) under four conditions, whose plausibility in this application is assessed as follows.

1. *Parallel trends.* In the absence of the reform, outcomes of treated and control parents would have evolved in parallel, conditional on covariates and fixed effects. The assumption concerns counterfactual trends and is not directly testable, but its pre-treatment analogue is: treated–control outcome gaps should be stable before 2009. This is examined graphically on unadjusted group means (Section 4.4), with regression-based event studies (Section 6.2), and with the pre-treatment placebo averages of the group-time estimator; the degree to which conclusions survive violations of the assumption is quantified in Section 7.4. As reported below, pre-trends are flat for benefit collection and less precisely estimated for unemployment duration.

2. *No anticipation.* Outcomes before the treatment date are unaffected by the forthcoming treatment. In this application the assumption is partially violated in one specific and observable way: the six-week interval between adoption and entry into force allowed households to time registration, and Appendix E documents excess claims in December 2008–January 2009. Part of this bunching may reflect ordinary intertemporal displacement of claims rather than anticipation of the reform itself; both interpretations imply the same remedy of excluding the adoption window. Claims filed earlier under the old rules retained their entitlement, so the violation operates through claim timing rather than through outcomes of already-granted claims; estimators sensitive to it (the group-time estimator) are applied excluding the adoption window, and the baseline estimates are shown

to be insensitive to this exclusion.

3. *No spillovers.* The reform must not affect the control group's outcomes. Treated spells constitute 0.15% of benefit spells in the sample, so general-equilibrium effects of the parental provision on other parents' outcomes are negligible by scale. A related compositional point is that the sample pools spells subject to different statutory county PBD categories; county fixed effects absorb the associated level differences, the statutory category enters the entitlement-based outcomes directly, and the stable-county sample of Appendix B removes the issue altogether. The concurrent change in the county threshold did affect the control group; this is a violation of a single-treatment framing rather than of the design, and it is addressed through county and time fixed effects and, directly, by re-estimation on counties whose PBD category never changed (Appendix B).

4. *Stable composition.* Because the design pools repeated cross-sections of claims, β compares different individuals before and after the reform. If the composition of the treated inflow changes at the reform date for reasons other than the reform, the comparison is contaminated. Treated eligibility requires an unemployed spouse who had not yet lost his or her entitlement, a condition whose incidence varies with labour-market conditions, so claimants entering treated status after 2008 may differ systematically from earlier entrants. This threat is addressed by covariate adjustment, by reweighting and matching (Section 5.3) and, most stringently, by individual fixed-effects specifications that compare the same person across regimes. The distinction matters for interpretation: the pooled specifications compare observationally similar – but not identical – individuals before and after the reform, whereas the fixed-effects specifications compare literally the same individuals observed under both regimes.

Because treatment timing is a single date common to all treated units, the two-way fixed-effects estimator does not suffer from the negative-weighting problems that arise under staggered adoption (Goodman-Bacon, 2021; Roth et al., 2023); estimators robust to treatment-effect heterogeneity are nevertheless reported as part of the robustness analysis.

5.2 Threats to identification

Four threats are addressed explicitly. *Concurrent regional reform.* The same amendment raised the county threshold from 125% to 150%, shortening the standard PBD from 12 to 6 months in reclassified counties. This affects the control group directly and the treated group only through the (now abolished) interaction of the parental rule with county categories. County fixed effects absorb time-invariant county differences, and the month fixed effects absorb the aggregate effect of reclassification; residual confounding would require differential sorting of treated parents across county categories coinciding with the reform. Appendix B restricts the sample to counties whose category did not change.

Common macroeconomic shocks. Aggregate labour-market conditions changed substantially over the analysis period. The design differences out shocks common to treated and control parents; the identifying assumption is violated only if such changes affected the two groups differentially, conditional on covariates and county. Event-study specifications and the formal sensitivity analysis of Rambachan and Roth (2023) assess this possibility.

Composition of the treated inflow. The incidence of the eligibility condition varies with labour-market conditions, and the annual treated inflow fell to its minimum around 2008–2009. Three responses are implemented: propensity-score reweighting and matching, which equalise observable composition between groups within periods; individual fixed-effects specifications, which identify β exclusively from within-person changes across repeat spells and are immune to selection on time-invariant unobservables; and group-specific linear trends, which replace the parallel-paths assumption with a parallel-growth assumption.

Strategic claim timing. The six-week interval between adoption (19 December 2008) and entry into force (1 February 2009) allowed forward-looking households to register under the old rules. Claim timing responses are documented in Appendix E; estimates excluding the months surrounding implementation are unchanged (Section 7.5).

5.3 Reweighting, matching and alternative estimators

Table 2 documents that treated and control parents differ in observable composition – education, urban residence, family structure and county type – and that these differences are not constant

across periods. Matching and reweighting re-establish comparability on observables before differencing. The propensity score – the probability of belonging to the treated group given covariates, $p(X) = \Pr(T = 1 \mid X)$ – suffices for this purpose: if treated and control units with the same X are comparable, they are also comparable conditional on $p(X)$ alone (Rosenbaum and Rubin, 1983). Two conditions are required: overlap, and – in the DiD context – conditional parallel trends, which need to hold only within covariate cells, a strictly weaker requirement than unconditional parallel trends (Heckman et al., 1997; Abadie, 2005). Overlap is verified in Appendix A and is favoured by the size of the comparison reservoir of about 650 control spells per treated spell.

The propensity score is estimated separately for the pre- and post-reform periods by a logit of treatment status on sex, a quadratic in age, work experience, education, urban residence, number of children, age of the youngest child, the relative county unemployment rate, occupation (one-digit groups), and benefit-start year indicators. Estimating by period ensures that treated claims are compared with control claims filed under the same regime and calendar conditions. Two estimators are used. First, inverse-probability weighting (IPW): specification (3) is estimated with weights

$$w_i \;=\; T_i + (1 - T_i)\,\frac{\hat{p}(X_i)}{1 - \hat{p}(X_i)}, \tag{4}$$

that is, unity for treated spells and the odds of treatment for control spells, restricted to the region of common support (Hirano et al., 2003). The weights reweight the control group to the covariate distribution of the treated, so that the weighted DiD recovers the ATT under conditional parallel trends (Abadie, 2005). Second, nearest-neighbour matching: each treated spell is matched, within period, to its three nearest control spells in terms of the estimated propensity score (caliper 0.05, matching with replacement), and specification (3) is re-estimated on the matched sample. The two estimators trade off differently between bias and variance – IPW retains all control observations, matching discards those far from the treated support – so agreement between them is informative.

As a further check, the group-time average treatment effect estimator of Callaway and Sant'Anna (2021) is applied with the doubly robust weighting of Sant'Anna and Zhao (2020), using benefit-start year as the time dimension, never-treated parents as the control group, and the

covariate set above. This estimator relaxes the functional-form assumptions of (3) and yields an event-study aggregation used for the sensitivity analysis of Rambachan and Roth (2023). These estimators do not all target identical estimands: the baseline, IPW-weighted and matched specifications estimate the ATT for the full treated inflow under progressively weaker comparability assumptions, the group-time estimator averages cohort-specific effects for the cohorts retained after excluding the adoption window, and the individual fixed-effects specification identifies the effect from the subpopulation of claimants with repeat spells. Agreement across them is therefore informative about robustness, but residual differences can reflect their different identifying populations as well as their different assumptions.

Finally, a within-group cohort design compares treated claims starting in February–July 2009 (the first six months of the 12-month regime) with treated claims starting in the same calendar months of 2008 (the old regime), using control parents in the same windows to difference out year effects; the design follows cohort-based comparisons exploiting entitlement variation across adjacent claim dates (Johnston and Mas, 2018). The identical placebo contrast for 2008 versus 2007, when no reform occurred, provides a specification test. Given the small treated inflow in these windows (24–76 claims per window), this design serves as a consistency check rather than a primary estimate.

6 Results

6.1 Main estimates

Table 3 reports estimates of β from (3). Each column corresponds to one outcome and each cell to a separate regression. The six rows correspond to estimators that rely on progressively different assumptions: the baseline two-way fixed-effects specification (parallel trends conditional on covariates); the same specification excluding the December 2008–January 2009 adoption window (robustness to claim timing); the IPW-weighted and nearest-neighbour matched specifications (parallel trends conditional on observable composition); the specification with a treated-group linear trend (parallel growth rather than parallel levels); and the individual fixed-effects specification (comparisons within the same person across regimes). Agreement across

rows indicates that an estimate is not an artefact of any single assumption. Each coefficient is expressed relative to the control group and net of common time effects: -2.92 in the first cell, for example, means that benefit collection of treated parents fell by 2.92 months more than that of control parents between the pre- and post-reform periods.

Table 3: Effect of the reduction in potential benefit duration: difference-in-differences estimates

	Benefit collection (months)	Unemployment duration (months)	Benefit exhausted (0/1)	Fraction of entitlement used	Exit to employment (0/1)
(1) Baseline TWFE	-2.92^{***}	-3.63^{***}	0.102^{***}	0.044^{***}	0.002
	(0.17)	(0.67)	(0.023)	(0.012)	(0.017)
(2) Excl. transition window	-2.99^{***}	-3.58^{***}	0.096^{***}	0.040^{***}	0.002
	(0.16)	(0.65)	(0.024)	(0.012)	(0.017)
(3) IPW (ATT weights)	-3.09^{***}	-3.81^{***}	0.109^{***}	0.049^{***}	0.002
	(0.16)	(0.65)	(0.022)	(0.012)	(0.017)
(4) PSM, NN(3)	-2.94^{***}	-3.89^{***}	0.132^{***}	0.063^{***}	0.001
	(0.18)	(0.71)	(0.027)	(0.015)	(0.021)
(5) Group-specific trend	-2.73^{***}	-4.03^{***}	0.102^{**}	0.061^{***}	-0.015
	(0.32)	(1.48)	(0.049)	(0.022)	(0.037)
(6) Individual FE	-2.31^{***}	-1.33	0.096^{**}	0.048^{**}	-0.017
	(0.26)	(0.86)	(0.042)	(0.022)	(0.033)
N (baseline)	1,767,335	1,767,335	1,767,335	1,767,335	1,767,335

Notes: Each cell reports the coefficient on the treated $\times$ post interaction from a separate regression. All specifications include covariates (sex, age, work experience, education, urban residence, number of children, age of youngest child), county fixed effects, and benefit-start month-by-year fixed effects. Row (2) excludes benefit starts in the December 2008–January 2009 adoption window; row (3) weights control spells by ATT weights from a period-specific propensity score on common support ($N = 1{,}693{,}832$); row (4) uses the nearest-neighbour matched sample ($N = 10{,}832$); row (5) adds a linear trend interacted with treatment status; row (6) includes individual fixed effects and is estimated on spells of individuals observed at least twice ($N = 1{,}009{,}166$). Standard errors in parentheses, clustered by county (329 clusters). TWFE = two-way fixed effects; IPW = inverse-probability weighting; PSM = propensity-score matching; NN(3) = matching on the three nearest neighbours. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Benefit collection duration falls by 2.92 months (s.e. 0.17) in the baseline specification. The estimate is -2.99 when the adoption window is excluded, -3.09 under IPW reweighting, -2.94 under nearest-neighbour matching, -2.73 with a treated-specific linear trend, and -2.31 with individual fixed effects; all are significant at the 1% level. Relative to the pre-reform treated mean of 14.7 months, the baseline estimate implies a reduction of 20%; scaled by the six-month reduction in entitlement, it corresponds to 0.49 months of benefit collection per month of PBD. This response can be read against the discontinuity evidence for the general population of Polish claimants: Jessen et al. (2026) find that extending PBD from 6 to 12 months raises the unemployment of directly affected recipients below the age of 50 by about 13%, and Jessen

et al. (2025) document that longer durations amplify the distortionary effect of higher benefit levels. The benefit-receipt response estimated here for parents sits within this pattern, while the absence of a detectable employment response, documented below, sets this category apart. The stability of the estimate across reweighting and matching indicates that it is not driven by differences in observable composition, and the within-individual estimate confirms the response holding time-invariant individual characteristics fixed.

Registered unemployment duration falls by 3.63 months (s.e. 0.67) in the baseline specification, with closely similar estimates under IPW (-3.81), matching (-3.89) and group trends (-4.03). The individual fixed-effects estimate, by contrast, is -1.33 months (s.e. 0.86) and not statistically distinguishable from zero. The divergence between the cross-sectional and within-individual estimates is consistent with a change in the composition of the treated inflow around the reform: individuals entering treated status after 2009 differ from earlier entrants in characteristics associated with shorter spells, and the cross-sectional estimator attributes part of this compositional shift to the reform. The within-individual estimate, identified from 1,021 ever-treated individuals with repeat spells, bounds the behavioural response of unemployment duration at a smaller and statistically insignificant magnitude – though this specification is also the most demanding of the data, an issue taken up in Section 8. Both estimates are reported throughout, and the discussion treats the cross-sectional estimate as an upper bound on the behavioural effect.

The probability of exhausting the entitlement increases by 10.2 percentage points (s.e. 2.3) in the baseline specification, from a pre-reform treated mean of 48.5%; estimates range from 9.6 to 13.2 points across specifications and remain significant with individual fixed effects. The fraction of entitlement used rises by 4.4 percentage points (4.0–6.3 across specifications). The shorter maximum duration thus binds more frequently: benefit receipt is compressed towards the new cap rather than terminating earlier of its own accord.

The probability of exiting the register to employment is unaffected. The baseline estimate is 0.002 (s.e. 0.017) against a pre-reform treated mean of 0.60, and no specification yields an estimate exceeding 0.02 in absolute value or approaching conventional significance levels. The 95% confidence interval of the baseline estimate excludes changes in the employment-exit prob-

ability larger than about ± 3.5 percentage points. A six-month reduction in entitlement therefore produced no detectable change in the probability that an affected parent left unemployment for a recorded job.

6.2 Parallel trends and event-study estimates

The validity of the design rests on the parallel-trends condition. The graphical assessment on unadjusted group means (Section 4.4) showed treated series fluctuating around a stable distance above the control series before 2009, with the sharp break at the reform date confined to the treated group. Figure 5 formalises that assessment with event-study estimates: coefficients on the interaction of treatment status with benefit-start year (base year 2008) from specification (3) with the interaction term replaced by the full set of treatment-by-year terms. The pre-2009 coefficients measure deviations of the treated–control gap from its 2008 level, conditional on covariates and fixed effects; parallel pre-trends require them to be jointly indistinguishable from a constant.

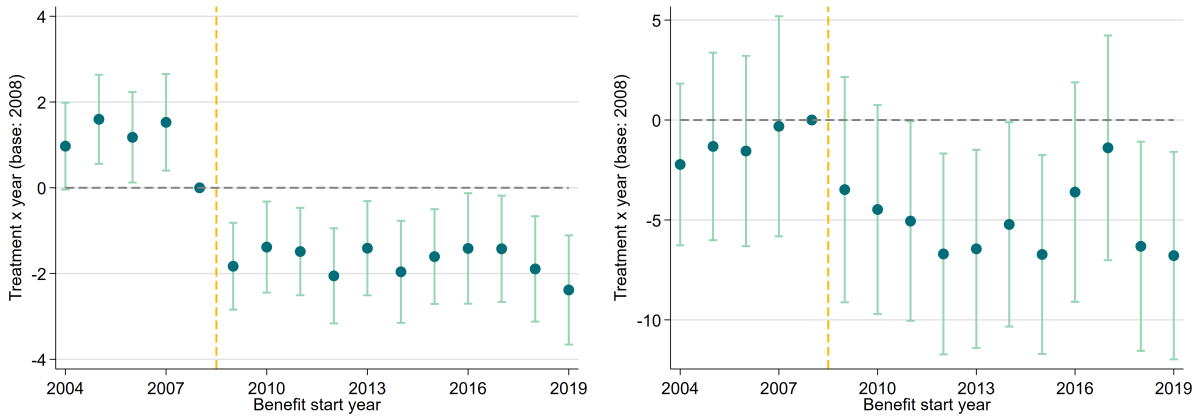

Figure 5: Event-study estimates: benefit collection and unemployment duration

Notes: Benefit collection duration (left) and registered unemployment duration (right). Treatment × benefit-start-year coefficients relative to 2008 from specification (3); 95% confidence intervals; standard errors clustered by county. Source: administrative records of Polish county labour offices; author's elaboration.

For benefit collection, the post-reform coefficients are uniformly negative (between -1.4 and -2.4 months relative to 2008) and the pre-reform coefficients fluctuate around $+1.3$ months without a trend. The level shift of the pre-reform coefficients relative to the 2008 base reflects the thinness of the base year: only 74 treated claims began in 2008, the lowest annual treated inflow in the sample, so the base-year group mean is imprecisely estimated. Re-basing against

the pooled 2004–2007 period, the pre-reform profile is flat and the post-reform drop is approximately -3 months, consistent with Table 3. For unemployment duration the pre-reform coefficients decline in absolute value towards 2008 (from -1.8 to -0.3) with wide confidence intervals, and the post-reform coefficients range from -1.4 to -6.8, with all but one year (2017) below -3.4. The event-study estimates from the IPW-weighted specification display the same pattern (Appendix D).

The group-time average treatment effects of Callaway and Sant'Anna (2021), which avoid the thin-base-year problem by weighting group-time contrasts and adjust for covariates by doubly robust estimation (Sant'Anna and Zhao, 2020), yield an average post-reform effect on benefit collection of -3.4 months (s.e. 0.6) with an average pre-reform placebo effect of 0.03 (s.e. 0.16); the hypothesis of parallel pre-trends conditional on covariates is not rejected. Because the estimator is defined under a no-anticipation condition, and Appendix E documents strategic timing of claims between the adoption and the implementation of the reform, spells with benefit start between August 2008 and January 2009 are excluded from this estimation; retaining them contaminates the 2008 base period with strategically timed claims and attenuates the estimates. For unemployment duration the corresponding average effect is negative and significant but larger in magnitude than the TWFE estimate, reflecting the long-difference structure of the estimator and the imprecision of the 2008 base; the sensitivity analysis in Section 7.4 therefore focuses on robust confidence sets rather than point magnitudes. The group-time event-study profiles are shown in Figure 6.

6.3 Exit destinations

Table 4 decomposes the effect on the composition of exits from the register. Before the reform, 60.1% of completed treated spells ended in employment (including business start-up), 31.8% in removal for non-compliance or withdrawal ("inactivity"), 7.1% in entry into active labour-market programmes, and 1.1% in retirement or pension transitions.

The reform left the destination structure essentially unchanged. The point estimates for employment ($+0.2$ p.p.) and inactivity ($+0.2$ p.p.) are close to zero and insignificant. The only statistically significant shift is a decline of 2.1 percentage points (s.e. 0.9) in exits through

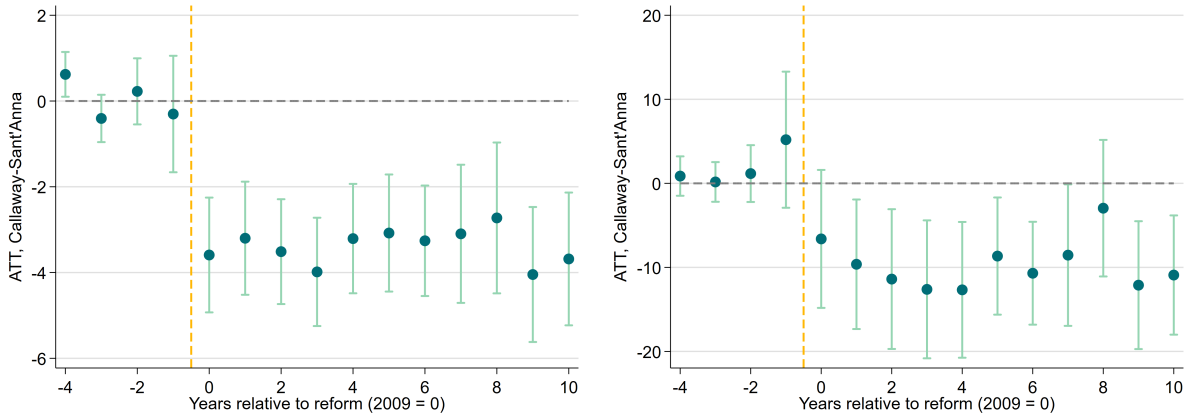

Figure 6: Group-time event-study estimates: benefit collection and unemployment duration

Notes: Group-time event-study estimates (Callaway and Sant'Anna, 2021), benefit collection duration (left) and registered unemployment duration (right). Doubly robust estimation with covariates; spells with benefit start August 2008–January 2009 excluded; 95% confidence intervals; standard errors clustered by county. Source: administrative records of Polish county labour offices; author's elaboration.

Table 4: Effect of the reform on exit destinations

Exit destination	DiD estimate	(SE)	Pre-reform mean, treated
Employment (incl. self-employment)	0.002	(0.017)	0.601
ALMP participation	-0.021^{**}	(0.009)	0.071
Retirement / pension	0.003	(0.004)	0.011
Inactivity	0.002	(0.018)	0.318
Censored / unknown	0.013^{*}	(0.007)	0.000

Notes: Linear probability models as in specification (3), one regression per destination. Destinations: employment (job take-up, business start-up); ALMP (intervention works, wage subsidies, programme participation); retirement/pension (retirement, disability pension, pre-retirement benefit, retirement age); inactivity (failure to confirm readiness to work); censored/unknown. Standard errors clustered by county. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

active labour-market programmes, quantitatively minor relative to the changes in duration outcomes. The small positive coefficient on the censored/unknown category (+1.3 p.p. against a pre-reform mean of zero) is mechanical rather than behavioural: spells that began late in the post-reform period are more likely to be still open at the end of the observation window, so recorded censoring rises for the late cohorts. The shortening of the benefit period thus compressed the time profile of benefit receipt without redirecting exits across recorded destinations.

These estimates inherit the measurement property noted in Section 4.2: destinations describe the reason recorded at deregistration, not observed subsequent activity. An exit recorded as employment is not verified against employer or social-insurance records, and an exit recorded as failure to confirm readiness to work may conceal unregistered employment as well as genuine withdrawal from the labour force. The finding of an unchanged destination structure should

accordingly be read as evidence that the reform did not alter what claimants reported upon leaving the register; whether the underlying post-exit activities changed cannot be established with these data.

7 Robustness and heterogeneity

This section examines, one by one, the conditions on which the main estimates rest: covariate balance and common support for the reweighted and matched estimators (Section 7.1, with diagnostics in Appendix A); the concurrent regional reform (Appendix B); the dependence of inference on asymptotic approximations (Section 7.3); the sensitivity of conclusions to violations of parallel trends (Section 7.4); and claim timing around the reform, assessed through an adjacent-cohort comparison with a placebo counterpart (Section 7.5) and through the inflow analysis of Appendix E. Section 7.6 reports heterogeneity.

7.1 Covariate balance and common support

The matched and reweighted estimators are informative only if the propensity score succeeds in equalising the observable composition of the two groups. Appendix A reports the standard diagnostics. In the raw data, imbalances are moderate in the pre-period: treated parents are less urban (standardised difference -0.191), have younger children (-0.191), more basic vocational education (0.130) and less tertiary education (-0.157). The imbalances change character in the post-period, where the largest by a wide margin is the relative county unemployment rate (-0.582): after the threshold change, post-reform treated claims are concentrated in counties with low relative unemployment, so an unadjusted comparison would partly contrast treated parents in slack pre-reform labour markets with control parents in tighter post-reform ones. After reweighting, every standardised difference in both periods falls below 0.024 in absolute value, and most fall below 0.01. The control distribution of the propensity score covers the treated distribution over its entire range in both periods; enforcing common support removes about 4% of control spells and none of the treated. The pseudo-R^2 of the propensity model is 0.044 in the pre-period and 0.028 in the post-period: treated status is only weakly predictable

from observables, as expected, because eligibility is determined primarily by the spouse's benefit history, which the covariate set does not contain. The weak predictability is not a deficiency of the model; it implies extensive overlap and stable weights, and it locates the relevant selection concern in unobservables – which the individual fixed-effects estimator addresses separately.

The substantive conclusion of the diagnostics is carried by the estimates themselves: reweighting and matching move the benefit-collection estimate from -2.92 to -3.09 and -2.94 respectively, and the unemployment-duration estimate from -3.63 to -3.81 and -3.89 (Table 3). Equalising observable composition, including the county unemployment-rate imbalance, leaves the estimates within one standard error of the baseline. Whatever part of the raw contrast was attributable to observable composition, it was small.

7.2 The concurrent regional reform

The 2009 amendment changed two rules at once: it reduced the parental entitlement and raised the county threshold governing standard PBD, which shortened the control group's entitlement from 12 to 6 months in reclassified counties. Appendix B re-estimates the baseline on the subsample of counties whose statutory category is identical under the 2008 and 2010 classifications – counties never reclassified, in which the control group's entitlement was constant throughout (85% of the estimation sample). Every estimate lies within one standard error of its baseline counterpart; the benefit-collection effect, for example, is -2.83 (s.e. 0.18) against -2.92 (0.17) in the full sample. The concurrent regional reform does not drive the estimated effect of the parental entitlement reduction.

7.3 Inference

The reported standard errors are clustered at the county level to allow arbitrary correlation of errors within local labour markets, both across individuals and over time. Cluster-robust inference is asymptotic in the number of clusters; although 329 clusters is a large number by conventional standards, the treated observations are few (2,722) and unevenly distributed across counties, a configuration in which the asymptotic approximation can understate uncertainty about interaction coefficients (Cameron et al., 2008). As a check, the main duration estimates

are re-tested with the wild cluster bootstrap, which resamples cluster-level residuals under the null hypothesis (Rademacher weights, 999 replications) and inverts the bootstrap test to obtain confidence sets; the procedure has good finite-cluster properties precisely in settings of this kind (Cameron et al., 2008; Roodman et al., 2019).

The bootstrap yields $p < 0.001$ for both duration outcomes. The bootstrap 95% confidence set is $[-3.27, -2.59]$ for the benefit-collection effect and $[-4.99, -2.25]$ for the unemployment-duration effect – in both cases nearly coincident with the analytic cluster-robust intervals implied by Table 3 ($[-3.25, -2.59]$ and $[-4.95, -2.31]$, respectively). The statistical significance of the main estimates is therefore not an artefact of the asymptotic variance approximation or of the uneven distribution of treated spells across clusters.

7.4 Sensitivity to violations of parallel trends

Flat pre-trends do not by themselves establish that trends would have remained parallel after 2009: pre-trend tests have limited power, and post-2009 labour-market conditions could in principle have affected the two groups differentially in a way that no pre-period observation can reveal. Rather than treating parallel trends as an all-or-nothing hypothesis, the approach of Rambachan and Roth (2023) asks how large a post-reform violation of parallel trends would have to be, relative to the violations actually observed before the reform, for the conclusion to change. Formally, the post-treatment differential trend is allowed to be as large as M times the largest deviation observed across consecutive pre-treatment periods, and a confidence set for the treatment effect is constructed that is valid for all violations within this bound. $M = 0$ corresponds to exact parallel trends; $M = 1$ allows post-reform confounding as large as anything seen in the pre-period; the "breakdown value" is the largest M at which the confidence set still excludes zero.

The procedure is applied to the event-study aggregation of the group-time estimator, using the four pre-reform event coefficients (2005–2008) as the benchmark for observed violations and the first post-reform coefficient as the target. For benefit collection, the robust confidence set for the first post-reform effect is $[-5.33, -1.62]$ at $M = 0.5$ and $[-5.95, -0.70]$ at $M = 1$, and first includes zero at $M = 1.5$ ($[-6.67, 0.28]$): the breakdown value lies between 1 and

1.5. The conclusion that the reform reduced benefit collection would therefore survive a post-reform confounding trend as large as the largest fluctuation observed in the four years before the reform, and would fail only under confounding half as large again. The flat pre-period profile (average pre-reform placebo effect of 0.03 months) makes a violation of this size difficult to motivate. For unemployment duration, by contrast, the original confidence interval of the first post-reform coefficient already includes zero, so the estimate is not robust at any M. This is consistent with the pattern of Section 6.1: the evidence for a causal reduction in unemployment duration is weaker on every dimension – within-individual identification, first-year timing and sensitivity analysis – than the evidence for the benefit-collection response.

7.5 Adjacent-cohort comparison, placebo and claim timing

The main design pools sixteen years of claims and relies on parallel trends over that horizon. A complementary design discards the long panel and compares only claims filed immediately around the reform: treated parents whose benefit started in February–July 2009 against treated parents whose benefit started in the same calendar months of 2008, with control parents claiming in the same two windows providing the second difference. Using identical calendar months in adjacent years removes seasonal composition by construction, and the design trades the long-horizon parallel-trends assumption for a much weaker local one – no treated-specific shock coinciding with the 2009 window – at the cost of statistical power, since only 24 treated claims fall in the 2008 window and 67 in the 2009 window. The identical placebo contrast one year earlier (February–July 2008 versus 2007, 76 treated claims in the 2007 window), when no reform occurred, provides a specification test. Table 5 reports both.

Table 5: Adjacent-cohort estimates and placebo

	Benefit collection	Unemp. duration	Benefit exhausted	Fraction used	Exit to employment
Feb–Jul 2009 vs. Feb–Jul 2008	-2.36^{**}	-7.85^{*}	0.38^{***}	0.13^{**}	0.02
	(1.00)	(4.24)	(0.09)	(0.06)	(0.11)
Placebo: Feb–Jul 2008 vs. 2007	-1.50	1.62	-0.16	-0.07	0.02
	(1.09)	(4.24)	(0.10)	(0.06)	(0.11)

Notes: Difference-in-differences between treated and control parents across adjacent same-calendar-month claim windows (February–July). Treated claims per window: 76 (2007), 24 (2008), 67 (2009). Covariates as in the baseline specification; standard errors clustered by county. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The 2009-versus-2008 contrast reproduces the main findings at the reform boundary. Benefit collection falls by 2.36 months (s.e. 1.00), within one standard error of the baseline estimate. The probability of exhaustion rises by 38 percentage points (s.e. 9.5), much more than the 10-point average effect of Table 3, as expected: the local contrast compares claimants facing the 12-month cap directly with claimants of the immediately preceding cohort facing the 18-month cap, whereas the pooled estimate averages the exhaustion response over a post-reform decade in which spells, counterfactual behaviour and labour-market conditions varied. The point estimate for unemployment duration is negative and large (-7.85, s.e. 4.24) but imprecise, and the employment-exit estimate is again indistinguishable from zero; with 24–67 treated claims per window, only large effects are detectable on these outcomes. In the placebo contrast every estimate is statistically insignificant, and the exhaustion and fraction-of-entitlement estimates reverse sign. The absence of placebo effects indicates that the 2009-versus-2008 differences are not an artefact of ordinary year-to-year variation in the treated group.

A final check concerns the transition period itself. Between the adoption of the amendment and its entry into force, and in the months immediately preceding adoption, claim timing was plausibly strategic (Appendix E), and the classification of claims into regimes is least sharp there. Excluding from the main specification all spells with benefit start between August 2008 and January 2009 leaves every estimate in Table 3 essentially unchanged; the benefit-collection estimate, for example, becomes -2.99 (s.e. 0.16), against -2.92 (0.17) in the baseline. Appendix E documents the underlying inflow behaviour: claims were advanced into the final weeks of the old regime, consistent with awareness of the forthcoming change; the observable composition of claims filed just before and just after implementation is balanced; and treated inflows did not decline after the reform.

7.6 Heterogeneity

Table 6 reports baseline estimates by subgroup for the three outcomes with non-zero average effects.

Four patterns are noteworthy. First, the reduction in benefit collection is somewhat larger for men (-3.32) than for women (-2.64), while the increase in exhaustion probability is con-

Table 6: Heterogeneity of the main estimates

Subsample	Benefit collection (months)	Unemployment duration (months)	Benefit exhausted (0/1)
By sex			
Men	-3.32^{***} (0.21)	-3.64^{***} (0.69)	0.042 (0.032)
Women	-2.64^{***} (0.21)	-3.84^{***} (1.09)	0.153^{***} (0.030)
By education			
Below tertiary	-3.05^{***} (0.16)	-4.04^{***} (0.69)	0.090^{***} (0.024)
Tertiary	-2.27^{***} (0.66)	0.41 (2.44)	0.173^{**} (0.079)
By place of residence			
Rural	-3.30^{***} (0.20)	-4.88^{***} (0.82)	0.076^{**} (0.031)
Urban	-2.72^{***} (0.21)	-2.56^{**} (1.03)	0.120^{***} (0.031)
By standard county PBD			
6-month counties	-3.19^{***} (0.19)	-3.25^{***} (0.82)	0.211^{***} (0.024)
12-month counties	-4.66^{***} (0.44)	-8.09^{***} (1.40)	0.079 (0.088)

Notes: Each cell reports the treated $\times$ post coefficient from the baseline specification (3) estimated separately on the indicated subsample. "Standard county PBD" refers to the 6- or 12-month PBD category of the claimant's county in the year the benefit was granted. Standard errors in parentheses, clustered by county. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

centrated among women (+15.3 p.p., against an insignificant +4.2 p.p. for men): after the reform, treated women are markedly more likely to draw the entitlement to its end. Second, the decline in unemployment duration is driven entirely by claimants without tertiary education (-4.04 versus an insignificant $+0.41$ for tertiary graduates), consistent with the concentration of benefit-dependent spells in the lower education groups. Third, effects on unemployment duration are larger in rural areas (-4.88 versus -2.56 urban). Fourth, effects are substantially larger in high-unemployment counties with a 12-month standard PBD (-4.66 on benefit collection and -8.09 on unemployment duration, against -3.19 and -3.25 in 6-month counties). In 12-month counties the reform eliminated the entire six-month advantage of treated parents over the standard entitlement, whereas in 6-month counties a six-month advantage remained; the estimated gradient is consistent with the size of the entitlement change relative to the local counterfactual. The exhaustion response displays the opposite gradient (+21.1 p.p. in 6-month counties versus +7.9 in 12-month counties), reflecting the higher post-reform bindingness of the cap where spells are long relative to the local benchmark.

8 Discussion

8.1 Interpretation

The estimates document a large and robust response of benefit receipt to the entitlement reduction, and an absence of any detectable response on the employment margin. Three elements of this pattern merit comment.

First, the compression of benefit collection (-2.9 to -3.1 months for a six-month entitlement cut), combined with the increase in exhaustion ($+10$ p.p.) and in the fraction of entitlement used ($+4$ p.p.), indicates that the previous 18-month entitlement was substantially utilised. Before the reform, treated parents drew the benefit for 14.7 of the 18 available months on average, the median duration of receipt was 17.2 months, the mean fraction of entitlement used was 0.83, and 49% of spells reached exhaustion. Treated parents did not, on average, leave benefits long before exhaustion under either regime. The shift of the exhaustion mass point from 18 to 12 months (Figure 1) and the reallocation of unemployment exits into the 12–18 month interval are the distributional counterparts of this compression. Part of the measured reduction is therefore mechanical – spells that would have continued beyond 12 months were truncated by statute – and the estimates should not be read as a pure behavioural response.

Second, the wedge between the cross-sectional and within-individual estimates for unemployment duration (-3.6 versus an insignificant -1.3) implies that duration estimates obtained from repeated cross-sections of claimants around this reform partly reflect composition. The eligibility condition ties treated status to the spouse's labour-market state, so the composition of the treated inflow co-moves with labour-market conditions – a co-movement documented directly in Section 4.1, where new treated claims fall from over 900 in 2004 to 74 in 2008 and recover after 2009. The post-2009 inflow differs from the earlier inflow in ways only partially captured by observables. The within-individual estimate suggests a behavioural response concentrated in benefit receipt itself, with at most a modest effect on total registered unemployment duration. Two qualifications apply. The fixed-effects specification is also the most demanding of the data: it draws only on the 1,021 ever-treated individuals with repeat spells and carries a standard error about twice that of the baseline, so limited power contributes to the attenua-

tion; and claimants with repeat spells need not be representative of all treated claimants. More generally, the wedge carries a methodological lesson: where eligibility for a group-specific entitlement depends on household labour-market states, the set of eligible claimants changes with labour-market conditions, and evaluation designs that pool repeated cross-sections without addressing this compositional change will attribute part of it to the policy.

Third, the zero effect on recorded exits to employment is estimated with enough precision to rule out changes larger than about ± 3.5 percentage points, bearing in mind that this indicator captures the reason declared at deregistration rather than verified post-exit employment. This distinguishes the group from populations studied elsewhere, for which duration reductions typically raise job-finding hazards (Filges et al., 2018; Schmieder and von Wachter, 2016), and from the general population of Polish claimants, for which PBD changes move unemployment itself (Jessen et al., 2026). The affected households consist, by construction, of two non-employed adults with dependent children. For such households the constraint on re-employment is unlikely to be the reservation-wage margin targeted by benefit-duration incentives: caregiving obligations restrict feasible working arrangements. The combination of benefit receipt compressing to the new cap with unresponsive employment exits is consistent with the interpretation that, for this group, the reform operated primarily as a reduction in transfer duration rather than as an activation instrument – an interpretation, it should be stressed, rather than a directly established mechanism. The magnitude of the duration response per month of entitlement (0.5 months of benefit collection; 0.2–0.6 months of unemployment duration depending on estimator) lies within the range reported in the survey literature (Lopes, 2022; Cohen and Ganong, 2024), at its upper end for the cross-sectional estimates and near its centre for the within-individual estimates.

8.2 Limitations

Several limitations qualify these conclusions. The register does not link spouses, so household-level outcomes – total household benefit receipt, the interaction of the two spouses' spells, intra-household substitution of job search – cannot be observed, although the studied entitlement is defined at the household level. Relatedly, the register contains no unique child identifier: chil-

dren are observed only through the birth dates of household members reported by each parent, so households cannot be merged precisely. Post-unemployment job quality, wages and employment stability are not recorded, and exit destinations are recorded at deregistration without verification against subsequent employment or social-insurance records. The absence of an effect on the employment-exit margin is therefore a statement about recorded exits rather than about verified post-exit employment, and it cannot be extended to statements about match quality (cf. Nekoei and Weber, 2017; Tatsiramos and van Ours, 2014). Exits recorded as removal for non-compliance cannot be separated into genuine inactivity and unregistered employment, a caveat common to register-based studies (Card et al., 2007a); the finding that the destination composition of exits is unchanged mitigates, but does not eliminate, this concern. The treated population is small and specific, and the estimates are local to it; they do not generalise to the population of benefit recipients at large. Finally, although the control group absorbs common shocks, the thin 2008 treated inflow means that the parallel-trends assumption cannot be verified in the immediate pre-reform year with precision; the sensitivity analysis bounds, rather than eliminates, the resulting uncertainty for the unemployment-duration outcome.

9 Conclusions

This paper has examined the causal effects of shortening potential benefit duration for a specific category of unemployment benefit recipients: parents of a child below the age of 15 whose spouse was also unemployed and had not yet lost benefit entitlement when the claimant's own entitlement began. The 2009 amendment to the Polish employment promotion law reduced this group's entitlement from 18 to 12 months while leaving all other parameters of their benefits unchanged – a discrete change in a single design margin of unemployment insurance, applied to households with no employed adult. I used administrative records of 1.77 million benefit spells of parents, of which 2,723 belong to the treated category, and estimated the reform's effects by difference-in-differences against other benefit-receiving parents of young children, supported by reweighting and matching, individual fixed effects, group-specific trends, a doubly robust group-time estimator, formal sensitivity analysis and an adjacent-cohort design with placebo tests.

Three sets of findings emerge. First, the entitlement reduction compressed benefit receipt: treated parents collected benefits for approximately three months less than the control-adjusted counterfactual implies – a reduction of about one fifth relative to their pre-reform mean, or roughly half a month per month of entitlement removed. The estimate is stable across all estimators (between -2.3 and -3.1 months), survives wild-bootstrap inference and sensitivity analysis with respect to parallel-trends violations, and is visible in the raw data as a displacement of the exhaustion mass point from 18 to 12 months. Second, the incidence of benefit exhaustion rose by about 10 percentage points and the fraction of entitlement used by about 4 percentage points: the shorter maximum binds more often, and treated parents draw a larger share of a smaller maximum. Third, the probability that a spell ends with a recorded exit to employment did not change – the estimate is close to zero in every specification – and the recorded destination structure of exits is otherwise essentially unchanged. Effects on total registered unemployment duration are the least conclusive: cross-sectional estimators indicate a decline of about 3.6 months, but the estimate falls to an insignificant -1.3 months under within-individual identification, a difference attributable to the time-varying composition of the treated inflow.

The findings carry three implications. First, in a group-targeted scheme, the length of potential benefit duration directly governed the amount of benefit actually received: the response of benefit collection was large, immediate and persistent across the post-reform decade. Second, in this population the duration margin did not operate as an activation instrument. The mechanism emphasised in the job-search literature – a shorter entitlement raising exit rates to employment near exhaustion – finds no support in the recorded employment exits of treated parents; their household situation plausibly places binding constraints on job take-up that are not relaxed by shortening income support. The reform's first-order effect on the affected households was a reduction in the duration of transfer income; whether it also changed household welfare, consumption or poverty cannot be determined from register data and is not claimed here. Third, the analysis illustrates a methodological point for the evaluation of group-targeted entitlements generally: where eligibility depends on household labour-market states, the composition of the eligible inflow co-moves with labour-market conditions, and estimates that do not address composition can misattribute compositional change to behavioural response.

The results speak to current policy. The 2025 Act on the Labour Market and Employment Services retained a 12-month entitlement for claimants in analogous family circumstances – including parents with an unemployed spouse who has not yet lost benefit entitlement, and single parents – while abolishing the geographic differentiation of benefit duration (Labour Market and Employment Services Act, 2025). The present results suggest that the generosity of such extensions determines primarily the duration of income support received by eligible families; adjustments to their length should not be expected, by themselves, to change re-employment outcomes for this population, and any activation objective for households with two non-employed parents would need to operate through instruments addressed to their specific constraints – childcare availability foremost among them – rather than through the length of the benefit period. Several extensions remain open. Linking spouses within the register would permit analysis of the household as the unit of treatment; linking register exits to social-insurance employment records would replace declared exit reasons with verified employment states and permit analysis of job quality after unemployment – the margin on which the welfare assessment of duration reductions ultimately depends; and the rule changes introduced in 2025 will, in time, generate further rule-based variation against which the present estimates can be tested out of sample.

List of abbreviations

ALMP: active labour market programme; ATT: average treatment effect on the treated; DiD: difference-in-differences; FE: fixed effects; IPW: inverse-probability weighting; NN: nearest neighbour; PBD: potential benefit duration; p.p.: percentage points; PSM: propensity-score matching; s.e.: standard error; TWFE: two-way fixed effects; UI: unemployment insurance.

Declarations

Availability of data and materials

The individual-level administrative records analysed in this study originate from the registers of the Polish public employment services (county labour offices) and contain confidential per-

sonal information; they cannot be made publicly available. The data were obtained and used within the SONATA BIS 8 project of the National Science Centre, Poland (grant no. DEC-2018/30/E/HS4/00335), under conditions set by the data provider; the data are private and cannot be shared by the author. Replication code for all estimates is available from the corresponding author on request.

Competing interests

The author declares no competing interests.

Funding

This work was supported by the RID START grant "Assessing the impact of reduced unemployment benefit duration on parental employment outcomes in Poland".

Authors' contributions

AS conceived the study, performed the analysis and wrote the manuscript. The author read and approved the final manuscript.

Acknowledgements

This article is based on a chapter of the author's doctoral dissertation, *Effects of Selected Labour Market Policies in Poland: A Quasi-Experimental Approach* (SGH Warsaw School of Economics, 2026), prepared under the supervision of Ewa Gałecka-Burdziak.

References

Alberto Abadie. Semiparametric difference-in-differences estimators. *Review of Economic Studies*, 72(1):1–19, 2005.

Daron Acemoglu and Robert Shimer. Efficient unemployment insurance. *Journal of Political Economy*, 107(5):893–928, 1999.

Act of 14 March 2014. Act of 14 March 2014 amending the Act on Employment Promotion and Labour Market Institutions and certain other acts. Journal of Laws (Dz.U.) 2014 item 598, 2014.

Act of 19 December 2008. Act of 19 December 2008 amending the Act on Employment Promotion and Labour Market Institutions and certain other acts. Journal of Laws (Dz.U.) 2009 no. 6 item 33, 2009. Entered into force in stages (1 Jan 2009, 1 Feb 2009, 1 Jul 2009, 1 Jan 2010); the time-degressive benefit profile applied from 1 January 2010.

Vera Adamchik. The effect of unemployment benefits on the probability of re-employment in Poland. *Oxford Bulletin of Economics and Statistics*, 61(1):95–108, 1999. doi: 10.1111/1468-0084.00118.

Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2009.

Orley Ashenfelter and David Card. Using the longitudinal structure of earnings to estimate the effect of training programs. *Review of Economics and Statistics*, 67(4):648–660, 1985.

Anthony B. Atkinson and John Micklewright. Unemployment compensation and labour market transitions: A critical review. *Journal of Economic Literature*, 29(4):1679–1727, 1991.

Martin Neil Baily. Some aspects of optimal unemployment insurance. *Journal of Public Economics*, 10(3):379–402, 1978.

Tito Boeri. *Structural Change, Welfare Systems, and Labour Reallocation: Lessons from the Transition of Formerly Planned Economies*. Oxford University Press, 2000.

Brantly Callaway and Pedro H. C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427, 2008.

David Card, Raj Chetty, and Andrea Weber. The spike at benefit exhaustion: Leaving the unemployment system or starting a new job? *American Economic Review*, 97(2):113–118, 2007a.

David Card, Raj Chetty, and Andrea Weber. Cash-on-hand and competing models of intertemporal behavior: New evidence from the labour market. *Quarterly Journal of Economics*, 122 (4):1511–1560, 2007b.

Raj Chetty. A general formula for the optimal level of social insurance. *Journal of Public Economics*, 90(10-11):1879–1901, 2006.

Raj Chetty. Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy*, 116(2):173–234, 2008.

Jonathan P. Cohen and Peter Ganong. Disemployment effects of unemployment insurance: A meta-analysis. NBER Working Paper 32832, National Bureau of Economic Research, 2024.

Employment Act. Act of 29 December 1989 on Employment. Journal of Laws (Dz.U.) 1989 no. 75 item 446, 1989.

Employment and Counteracting Unemployment Act. Act of 14 December 1994 on Employment and Counteracting Unemployment. Journal of Laws (Dz.U.) 1995 no. 1 item 1, 1994.

Employment and Unemployment Act. Act of 16 October 1991 on Employment and Unemployment. Journal of Laws (Dz.U.) 1991 no. 106 item 457, 1991.

Employment Promotion and Labour Market Institutions Act. Act of 20 April 2004 on Employment Promotion and Labour Market Institutions. Journal of Laws (Dz.U.) 2004 no. 99 item 1001, 2004.

Norman T. Feather. *The Psychological Impact of Unemployment*. Springer-Verlag, New York, 1990.

Trine Filges, Anders Bruun Jonassen, and Anne-Marie Klint Jørgensen. Reducing unemployment benefit duration to increase job finding rates: A systematic review. *Campbell Systematic Reviews*, 14(1):1–194, 2018. doi: 10.4073/csr.2018.2.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

James J. Heckman, Hidehiko Ichimura, and Petra E. Todd. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4):605–654, 1997.

Keisuke Hirano, Guido W. Imbens, and Geert Ridder. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

Jonas Jessen, Robin Jessen, Andrew C. Johnston, and Ewa Gałecka-Burdziak. Moral hazard among the employed: Evidence from regression discontinuity. NBER Working Paper 33450, National Bureau of Economic Research, 2025.

Jonas Jessen, Robin Jessen, Ewa Gałecka-Burdziak, Marek Góra, and Jochen Kluve. The micro and macro effects of changes in the potential benefit duration. *Review of Economic Studies*, 93(3):1926–1962, 2026. doi: 10.1093/restud/rdaf056.

Andrew C. Johnston and Alexandre Mas. Potential unemployment insurance duration and labor supply: The individual and market-level response to a benefit cut. *Journal of Political Economy*, 126(6):2480–2522, 2018.

Katarzyna Kalinowska. The stabilizing role of unemployment benefits in Poland. *Central European Review of Economics & Finance*, 12(2):21–40, 2016.

Edward L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.

Lawrence F. Katz and Bruce D. Meyer. The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of Public Economics*, 41(1):45–72, 1990.

Dorian Kessler and Debra Hevenstone. The impact of unemployment benefits on birth outcomes: Quasi-experimental evidence from European linked register data. *PLOS ONE*, 17(3): e0264544, 2022. doi: 10.1371/journal.pone.0264544.

Jonas Kolsrud, Camille Landais, Peter Nilsson, and Johannes Spinnewijn. The optimal timing of unemployment benefits: Theory and evidence from Sweden. *American Economic Review*, 108(4-5):985–1033, 2018. doi: 10.1257/aer.20160816.

Oliwia Komada, Paweł Strzelecki, and Joanna Tyrowicz. A regression discontinuity evaluation of reducing early retirement eligibility in Poland. *Labour Economics*, 58:131–142, 2019.

Alan B. Krueger and Bruce D. Meyer. Labour supply effects of social insurance. In Alan J. Auerbach and Martin Feldstein, editors, *Handbook of Public Economics*, volume 4, pages 2327–2392. Elsevier, 2002.

Labour Market and Employment Services Act. Act of 20 March 2025 on the Labour Market and Employment Services. Journal of Laws (Dz.U.) 2025 item 620, 2025. URL https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20250000620.

Marta Lachowska, Isaac Sorkin, and Stephen A. Woodbury. Employers and unemployment insurance take-up. *American Economic Review*, 115(8):2529–2573, 2025. doi: 10.1257/aer.20230195.

Rafael Lalive. Unemployment benefits, unemployment duration, and post-unemployment jobs: A regression discontinuity approach. *American Economic Review*, 97(2):108–112, 2007.

Rafael Lalive. How do extended benefits affect unemployment duration? A regression discontinuity approach. *Journal of Econometrics*, 142(2):785–806, 2008.

Rafael Lalive, Jan C. van Ours, and Josef Zweimüller. How changes in financial incentives affect the duration of unemployment. *Review of Economic Studies*, 73(4):1009–1038, 2006.

Camille Landais. Assessing the welfare effects of unemployment benefits using the regression kink design. *American Economic Journal: Economic Policy*, 7(4):243–278, 2015.

Thomas Le Barbanchon. The effect of the potential duration of unemployment benefits on unemployment exits to work and match quality in France. *Labour Economics*, 42:16–29, 2016.

Thomas Le Barbanchon, Johannes F. Schmieder, and Andrea Weber. Job search, unemployment insurance, and active labor market policies. In *Handbook of Labor Economics*, volume 5. Elsevier, 2024. URL https://www.nber.org/papers/w32720.

Lars Ljungqvist and Thomas J. Sargent. The European unemployment dilemma. *Journal of Political Economy*, 106(3):514–550, 1998.

Marta C. Lopes. A review on the elasticity of unemployment duration to the potential duration of unemployment benefits. *Journal of Economic Surveys*, 36(4):1212–1224, 2022.

Ramon Marimon and Fabrizio Zilibotti. Unemployment vs. mismatch of talents: reconsidering unemployment benefits. *Economic Journal*, 109(455):266–291, 1999.

Casey McQuillan and Brendan Moore. Barriers to benefits: Unemployment insurance take-up and labor market effects. Working paper, Washington Center for Equitable Growth, 2025.

Bruce D. Meyer. Unemployment insurance and unemployment spells. *Econometrica*, 58(4): 757–782, 1990.

Dale T. Mortensen. Unemployment insurance and job search decisions. *Industrial and Labour Relations Review*, 30(4):505–517, 1977.

Arash Nekoei and Andrea Weber. Does extending unemployment benefits improve job quality? *American Economic Review*, 107(2):527–561, 2017. doi: 10.1257/aer.20150528.

Stephen J. Nickell. The effect of unemployment and related benefits on the duration of unemployment. *Economic Journal*, 89(353):34–49, 1979.

OECD. Benefits and wages: Net replacement rates in unemployment. OECD Benefits and Wages indicators, 2024. URL https://www.oecd.org/social/benefits-and-wages/.

Patrick A. Puhani. Poland on the dole: The effect of reducing the unemployment benefit entitlement period during transition. *Journal of Population Economics*, 13(1):35–44, 2000. doi: 10.1007/s001480050121.

Ashesh Rambachan and Jonathan Roth. A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591, 2023.

Mari Rege, Kjetil Telle, and Mark Votruba. Parental job loss and children's school performance. *Review of Economic Studies*, 78(4):1462–1489, 2011. doi: 10.1093/restud/rdr002.

Knut Røed and Tao Zhang. Does unemployment compensation affect unemployment duration? *Economic Journal*, 113(484):190–206, 2003.

David Roodman, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. Fast and wild: Bootstrap inference in Stata using boottest. *Stata Journal*, 19(1):4–60, 2019.

Paul R. Rosenbaum and Donald B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

Jonathan Roth, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244, 2023.

Pedro H. C. Sant'Anna and Jun Zhao. Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122, 2020.

Johannes F. Schmieder and Till von Wachter. The effects of unemployment insurance benefits: New evidence and interpretation. *Annual Review of Economics*, 8:547–581, 2016. doi: 10.1146/annurev-economics-080614-115758.

Konstantinos Tatsiramos and Jan C. van Ours. Labour market effects of unemployment insurance design. *Journal of Economic Surveys*, 28(2):284–311, 2014. doi: 10.1111/joes.12005.

Joanna Tyrowicz and Piotr Wójcik. Regional dynamics of unemployment in Poland: A convergence approach. In *The Labour Market Impact of the EU Enlargement: A New Regional Geography of Europe?*, pages 149–173. Physica-Verlag, Heidelberg, 2010.

Appendix

A Covariate balance and common support

This appendix reports the diagnostics underlying the reweighted and matched estimators of Section 5.3. Balance is measured by standardised differences – the difference in group means divided by the square root of the average of the group variances – computed for each covariate separately by period, first in the raw data and then after applying the ATT weights (4). Unlike t-statistics, standardised differences do not mechanically grow with sample size, which matters here given 1.75 million control spells; values below 0.10 are conventionally taken to indicate adequate balance. Table A1 reports the results discussed in Section 7.1: after reweighting, every standardised difference in both periods falls below 0.024 in absolute value, and the county unemployment-rate imbalance falls from -0.582 to 0.024.

Table A1: Standardised differences before and after IPW reweighting

	Before Feb 2009		After Feb 2009	
	Raw	Weighted	Raw	Weighted
Female	−0.115	0.000	−0.022	0.000
Age at registration	0.031	0.000	−0.100	0.006
Work experience (years)	0.093	0.000	−0.028	0.002
Urban residence	−0.191	0.000	−0.027	0.001
Number of children	0.146	0.000	0.085	−0.001
Age of youngest child	−0.191	0.000	−0.246	0.007
Relative county unemployment rate	−0.131	0.000	−0.582	0.024
Education: primary or less	0.020	−0.001	−0.003	0.000
Education: basic vocational	0.130	0.000	0.001	−0.001
Education: secondary general	−0.049	0.000	0.008	−0.001
Education: secondary vocational	−0.007	0.001	0.060	−0.001
Education: post-secondary	−0.104	0.000	0.044	−0.001
Education: tertiary	−0.157	0.001	−0.105	0.004

Notes: Standardised differences defined as the difference in group means divided by the square root of the average of group variances. "Weighted" columns apply ATT weights from the period-specific propensity score, common support imposed.

Common support is examined in Figure A1, which plots kernel density estimates of the propensity score for treated and control spells, separately by period. The control distribution covers the treated distribution over its entire range in both periods: for every treated spell there exist control spells with a similar propensity score, which is the requirement for matching and

reweighting to be feasible without extrapolation. Enforcing common support – dropping control spells with scores outside the treated range within their period – removes about 4% of control spells and none of the treated.

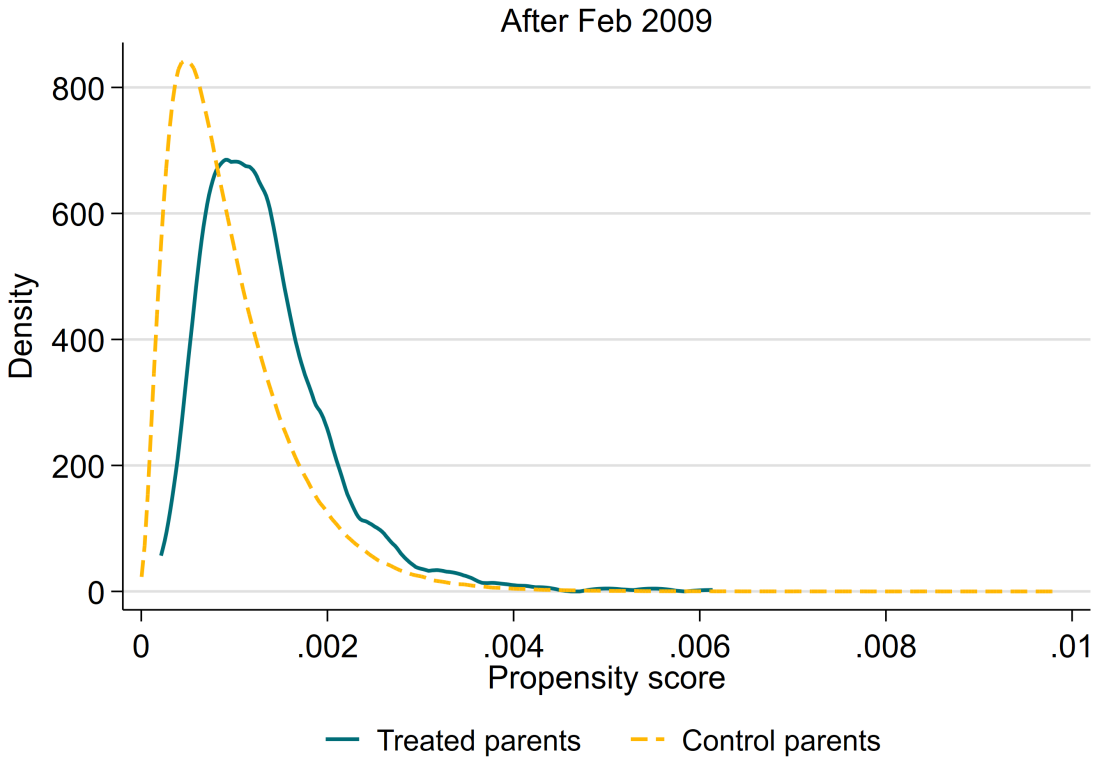

Figure A1: Propensity-score distributions of treated and control spells, by period
Notes: Kernel density estimates of the estimated propensity score for treated and control spells, before (left) and after (right) the reform. Source: administrative records of Polish county labour offices; author's elaboration.

B Stable-county robustness

The 2009 amendment changed two rules at once: it reduced the parental entitlement (the treatment of interest) and raised the county threshold from 125% to 150% of the national average unemployment rate. The second change shortened the standard entitlement from 12 to 6 months in every county whose relative unemployment rate fell between the two thresholds, and it affected the control group directly: control parents in reclassified counties experienced their own entitlement reduction at approximately the same date. If such counties were disproportionately represented among either group, the estimated treated $\times$ post coefficient could partly reflect this second reform rather than the parental provision.

The check removes the affected counties altogether. Each county is classified under the 2008 rule (12-month category if its relative rate on the 2008 reference date exceeded 125%) and under the 2010 rule (12-month category if its relative rate on the 2010 reference date exceeded 150%). The sample is then restricted to counties whose category is identical under both classifications: counties that were never reclassified and in which the control group's statutory entitlement was therefore constant throughout. The restriction retains $N = 1{,}503{,}400$ spells, 85% of the estimation sample. Table A2 reports the estimates.

Table A2: Estimates on the sample of counties never reclassified by the 2009 threshold change

Outcome	Estimate	(SE)
Benefit collection duration (months)	-2.83	(0.18)
Unemployment duration (months)	-3.73	(0.74)
Benefit exhausted (percentage points)	$+11.7$	(2.5)
Fraction of entitlement used (percentage points)	$+4.9$	(1.3)
Exit to employment (0/1)	0.0002	(0.017)

Notes: Baseline specification (3) estimated on the subsample of counties whose statutory PBD category (6 or 12 months) is identical under the 2008 classification (12-month category if the relative county unemployment rate exceeded 125% of the national average on the 2008 reference date) and under the 2010 classification (threshold of 150% on the 2010 reference date): $N = 1{,}503{,}400$ spells, 85% of the estimation sample. Standard errors clustered by county.

Each estimate is within one standard error of its baseline counterpart in Table 3. The concurrent regional reform, whatever its effects on the reclassified counties, does not drive the estimated effect of the parental entitlement reduction.

C Detailed descriptive statistics

Table A3 reports means, standard deviations and medians of the continuous covariates and outcomes of the analysis sample, by treatment group and reform period. The comparison of means and medians makes the skewness of the duration outcomes explicit. Mean unemployment duration among treated parents exceeds its median by more than four months in the pre-reform period, reflecting a right tail of very long spells, while the medians of benefit collection sit almost exactly at the statutory caps (17.2 months before the reform, 12.0 after).

Table A3: Detailed descriptive statistics by treatment group and reform period

	Control parents		Treated parents	
	Before Feb 2009	After Feb 2009	Before Feb 2009	After Feb 2009
Age at registration				
Mean	34.67	35.44	34.88	34.79
SD	(7.26)	(6.77)	(6.87)	(6.11)
Median	33.5	34.7	34.0	34.0
Work experience (years)				
Mean	10.25	9.35	10.90	9.18
SD	(7.10)	(6.36)	(7.01)	(5.99)
Median	8.5	7.9	9.1	8.2
Number of children				
Mean	1.70	1.79	1.83	1.87
SD	(0.87)	(0.91)	(0.91)	(0.87)
Median	2	2	2	2
Age of youngest child				
Mean	6.14	5.87	5.36	4.86
SD	(4.15)	(4.27)	(3.91)	(3.93)
Median	6	5	4	3
Benefit collection duration (months)				
Mean	5.92	5.24	14.74	10.42
SD	(3.84)	(3.50)	(4.12)	(2.26)
Median	5.98	5.59	17.15	12.03
Unemployment duration (months)				
Mean	11.73	11.22	24.06	19.40
SD	(14.24)	(13.42)	(18.66)	(14.43)
Median	7.85	7.10	19.65	14.95
Fraction of entitlement used				
Mean	0.67	0.66	0.83	0.87
SD	(0.37)	(0.39)	(0.23)	(0.19)
Median	0.76	0.67	0.97	1.00
Spells	580,631	1,185,413	1,598	1,125

Notes: Analysis sample as in Table 2. The fraction of entitlement used can marginally exceed one because recorded benefit days may exceed the statutory entitlement by a few days of administrative tolerance.

D Additional event-study results

Figure A2 reports the event-study estimates from the IPW-weighted specification, referred to in Section 6.2. The profiles mirror the unweighted estimates of Figure 5. Table A4 reports the full set of covariate coefficients underlying the baseline specification of Table 3.

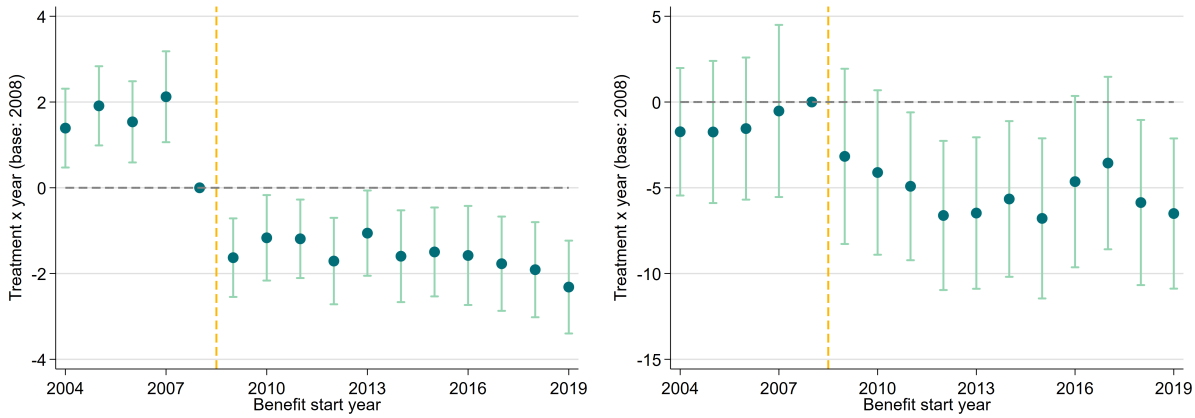

Figure A2: IPW-weighted event-study estimates: benefit collection and unemployment duration

Notes: IPW-weighted event-study estimates, benefit collection duration (left) and registered unemployment duration (right). Treatment × benefit-start-year coefficients relative to 2008; 95% confidence intervals; standard errors clustered by county. Source: administrative records of Polish county labour offices; author's elaboration.

Table A4: Full baseline regression estimates

	(1) Benefit	(2) Unemp.	(3) Exhausted	(4) Fraction	(5) Employment
Treated × post	−2.920***	−3.631***	0.102***	0.0437***	0.00201
	(0.166)	(0.672)	(0.0227)	(0.0117)	(0.0166)
Treated	8.674***	11.29***	0.0149	0.119***	−0.0576***
	(0.123)	(0.490)	(0.0178)	(0.0105)	(0.0128)
Female	1.227***	5.458***	0.186***	0.146***	−0.0769***
	(0.0361)	(0.144)	(0.00355)	(0.00278)	(0.00380)
Age at registration	0.0628***	0.228***	0.00828***	0.00562***	−0.00494***
	(0.00138)	(0.00564)	(0.000163)	(0.000138)	(0.000175)
Work experience (years)	−0.0116***	−0.0598***	−0.00405***	−0.00280***	0.00563***
	(0.00128)	(0.00464)	(0.000147)	(0.000120)	(0.000186)
Urban residence	−0.138***	−0.592***	−0.0129***	−0.0125***	0.00959***
	(0.0163)	(0.0532)	(0.00180)	(0.00155)	(0.00173)
Number of children	−0.0468***	0.100***	0.0000876	0.000734	−0.00159***
	(0.00553)	(0.0173)	(0.000613)	(0.000519)	(0.000597)
Age of youngest child	−0.101***	−0.469***	−0.0141***	−0.0110***	0.0106***
	(0.00243)	(0.00787)	(0.000199)	(0.000173)	(0.000203)
N	1,767,335	1,767,335	1,767,335	1,767,335	1,767,335

Notes: Baseline specification of Table 3, row (1). Education dummies, county fixed effects, and benefit-start month-by-year fixed effects included but not shown. Standard errors in parentheses, clustered by county. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Claim timing and inflow composition

This appendix documents behaviour on the entry margin around the 2009 reform. The patterns reported here are treated as descriptive context rather than as results in their own right: the monthly number of new treated claims is small – typically in the single digits – and county-level counts are smaller still, so that a single additional claim can move a local inflow by tens of percent. Claim inflows nevertheless matter for two reasons. As a threat to identification, strategic timing or selective entry around the reform date would change who is observed in each regime, in violation of the no-anticipation and stable-composition conditions of Section 5.1. As descriptive evidence, the literature documents that entitlement rules affect the timing of claims and separations (e.g. Filges et al., 2018; Jessen et al., 2025), so inflow responses indicate whether the affected households understood and reacted to the rules. Figure A3 plots monthly counts of new benefit claims for both groups over 2005–2012.

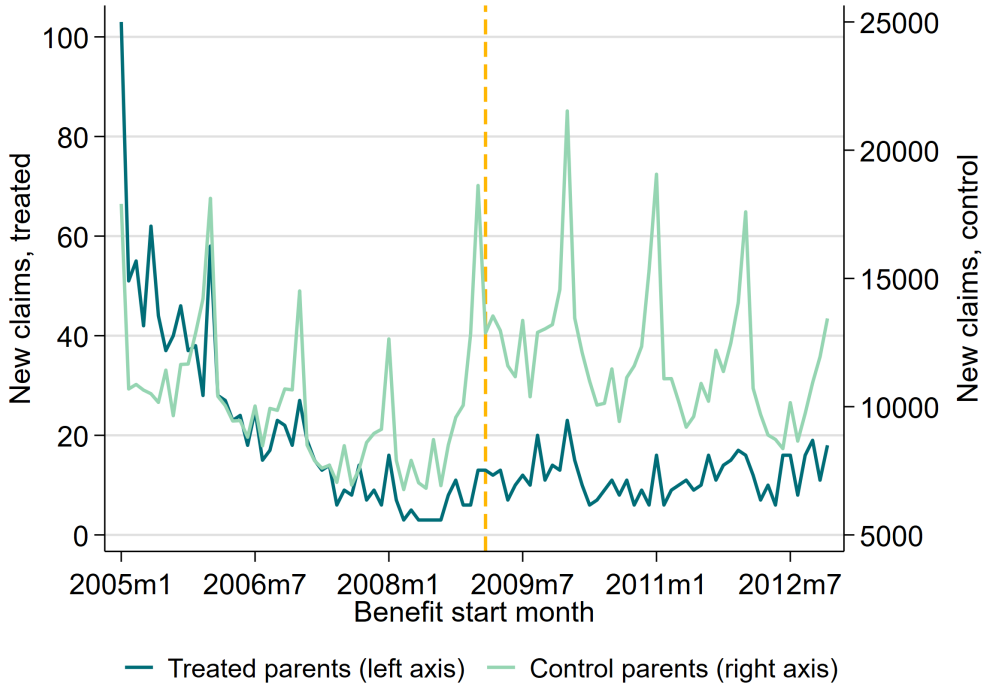

Figure A3: Monthly inflows of new benefit claims, treated and control parents, 2005–2012

Notes: Monthly counts of new benefit claims, treated parents (left axis) and control parents (right axis). The dashed vertical line marks the entry into force of the 2009 amendment. Source: administrative records of Polish county labour offices; author's elaboration.

Three features of the inflow series are examined. The first is claim timing around implementation. The amendment was adopted on 19 December 2008 and entered into force on 1 February

2009, and the transitional rule preserved the 18-month entitlement for benefits granted before the latter date; a household able to advance its registration by a few weeks could therefore secure six additional months of potential benefits. A Poisson regression of monthly treated claim counts (2005–2012) on a post-reform indicator, a linear trend, a trend break, calendar-month effects, and separate indicators for December 2008 and January 2009 quantifies the response. The January 2009 count is 98% above the level implied by trend and seasonality (incidence-rate ratio 1.98, $p < 0.001$), and the December 2008 count is 45% above it ($p = 0.004$). In absolute terms the excess is roughly a doubling of a small monthly flow (13 claims in January 2009 against a baseline of about 6) – precisely the order of magnitude at which such percentage changes should be read descriptively. The bunching indicates that affected households were informed about the forthcoming change and timed claims accordingly – behaviour analogous to the strategic timing of claims and separations documented around other entitlement changes.

The second feature is the composition of claims around the cutoff. If the households that accelerated their claims differed systematically from those that did not, the pre-reform window would be selectively composed and the regime comparison distorted. Comparing treated claims filed in the six months before implementation (August 2008–January 2009) with those filed in the six months after (February–July 2009), none of the individual characteristics differs significantly: sex, age, work experience, urban residence, number of children, age of the youngest child and tertiary education all yield p-values above 0.10. The single significant difference is the share of claims filed in 12-month counties (8.5% before against 31.3% after, $p = 0.003$), which does not indicate selection: county categories for 2009 claims are set by county unemployment rates measured in mid-2008, so rising local unemployment rates mechanically raised the share of high-unemployment counties among later claims. The share nevertheless remains below the 37.2% average for the full pre-reform period reported in Table 2, because the qualifying threshold itself rose from 125% to 150% in February 2009, which works in the opposite direction. Observable selection into the pre-reform window is therefore absent, and – as reported in Section 7.5 – excluding the entire window leaves the estimates essentially unchanged, so any selection on unobservables within these six months is inconsequential for the results.

The third feature is the level of inflows after the reform. A persistent decline in treated

claims relative to control claims would indicate that the shorter benefit duration discouraged take-up of the parental provision – an extensive-margin response that would itself be a policy-relevant effect and would change the composition of the post-reform treated group. The check regresses the difference in log monthly inflows (treated minus control) on a post-reform indicator, a trend and calendar-month effects. The post-reform coefficient is $+0.49$ (s.e. 0.22): relative inflows rose rather than fell. The increase is consistent with the mechanics of eligibility rather than with any effect of the entitlement change: qualifying for the parental extension requires a spouse who is unemployed and had not yet lost benefit entitlement when the claimant's began, a configuration whose incidence rises when unemployment is high, and the years after 2008 mechanically generated more eligible households. There is, accordingly, no evidence that the reform reduced take-up of the provision; the estimates of Section 6 describe the response of benefit receipt conditional on claiming, not a contraction of the claimant population.

In sum, the entry margin shows exactly one detectable behavioural response – the advancement of claims into the final weeks of the 18-month regime – and this response is localised, observable and inconsequential for the estimates once the surrounding months are excluded. Composition at the cutoff is balanced on observables, and take-up did not decline. Read descriptively, the entry margin neither threatens the internal validity of the main estimates nor reveals an additional margin of adjustment beyond claim timing.