Welfare-to-what ? Experimental evaluation of an activation programme for single mothers in poverty in France

Arthur Heim[1]

25 March, 2024, Preliminary version. Last update

[1] Research and evaluation officer, department of statistics, research and studies, National family allowance fund (Cnaf) PhD Candidate, Labour and public economics research unit - Paris School of Economics (EHESS). arthur.heim@cnaf.fr, arthur.heim@psemail.eu

Contents

Abstract

A quarter of households are single parents with their children, with mothers being the primary caregivers in the majority of these cases. This demographic represents a large share of the most vulnerable populations and is therefore the focus of anti-poverty policies, which now rely heavily on active active labour market policies. This study examines the causal effects of an intensive welfare-to-work programme targeting single parents on long-term welfare in France. In the experimental sample, 94% were women and 98% lived below the poverty line. The programme started in 2018 for three years and was extended for two more years because of the Covid-19 pandemic. It provided year-long intensive support, including individual and group sessions with highly trained social workers and childcare. The cost per participant was approximately € 2800, 4 times the usual spending for social support. Using a block-randomised encouragement design, 844 encouraged single-parents of 4 cohorts were invited to meetings presenting the programme, with 38% signing-up on average, whereas the 828 in the control group received no intervention. The analysis shows that participation increased from 28% to 48% from the 1st to 4th cohort, improved by more threatening invitation letters on the one hand, and more human enrollment through individual meetings in the programme's premises and testimonies from former participants on the other. Compliers are more likely to have less than a high school diploma, be among the poorest half, in their thirties, and registered with the Employment agency. The programme slows job finding rates during its first half, leading to a lock-in effect on poverty rates, disposable incomes, and employment. These effects swiftly fade-out and there are no average effect by the programme's end, for the following year, and up to 45 months since random assignment for the three first cohorts. Participants have higher employment rates than other comparison groups, but this difference is entirely driven by selection biais. The latter is so strong that estimates using the next-best identification strategy - modern doubly-robust difference-in-differences - fails to include experimental estimates in the confidence intervals. I provide additional analyses on distributional effect on incomes and treatment effect heterogeneity. These results are in line with the literature showing that welfare-to-work programmes and reforms' effects on impoverished single parents' employment are often weak, and strongly depend on a good buisness cycle. Yet to my knowledge, this is the only experimental evidence of the (absence) of benefits from welfare-to-work policies for poor single mothers in France. These results are therefore highly relevant to the political debates in France marked by a particularly violent conservative turn towards the most vulnerable.

- **JEL Classification Numbers** : I38, J16, J18
- **Keywords**: Welfare-to-Work, single parents, active labour market policy, long-term unemployment

Résumé

Les familles monoparentales représentent environ un quart des ménages avec enfants et sont principalement dirigées par des femmes. Elles font souvent face à d'importantes difficultés les plaçant au coeur des politiques de lutte contre la pauvreté, désormais largement centrées sur des politiques d'activation du marché du travail. Cette étude examine les effets causaux d'un programme d'accompagnement intensif visant à faciliter le retour à l'emploi de familles monoparentales, bénéficiaires longue durée des minima sociaux en France. Dans l'échantillon expérimental, 94 % sont des femmes et 98 % vivent en dessous du seuil de pauvreté. Le programme a débuté en 2018 pour trois ans et a été prolongé de deux années suite à la crise Covid-19. Il proposait un accompagnement intensif d'un an, comprenant des séances individuelles et de groupe avec des travailleurs sociaux hautement qualifiés et des possibilités de garde d'enfant. Le coût par participante était d'environ 2800 €, soit quatre fois les dépenses habituelles pour l'acompagnement social. En utilisant un recrutement basé sur un encouragement randomisé stratifié, 844 parents des 4 premières cohortes ont été invités à des présentations du programme, avec une adhésion moyenne de 38 %, tandis que les 828 des groupes témoins n'ont reçu aucune intervention. Les analyses montrent que la participation s'est accrue de 28 % à 48 % entre la première et quatrième cohorte, favorisée par un courrier plus menaçant d'un côté, et un accueil plus humain via des rendez-vous individuels sur les lieux de l'accompagnement et des témoignages de parents des cohortes précédentes de l'autre. Les participantes sont plus susceptibles d'être trentenaires, parmi les plus pauvres, les moins diplômées et déjà inscrites à pôle emploi. Dans les premiers 9 mois, le programme ralentit les reprises d'emploi, entraînant un effet de *lock-in* sur le taux de pauvreté, les revenus disponibles et l'emploi. Ces effets disparaissent d'ici la fin du programme, sans amélioration l'année suivante, et jusqu'à 45 mois depuis le tirage au sort pour les trois premières cohortes. Le taux d'emploi des participantes est pourtant plus élevé que celui de tous les groupes de comparaison. Mais cet écart relève uniquement du biais de sélection dû à des différences inobservables. Le biais de sélection est si fort que les estimations utilisant la meilleure stratégie alternative - des différences de différences appariées et doublement robustes - rejettent les estimations expérimentales des intervalles de confiance à 95 %. J'analyse également l'hétérogénéité de l'effet du programme et ceux sur la distribution des revenus potentiels. Au final ces résultats s'ajoutent à une littérature abondante et croissante pointant les effets faibles ou souvent délétères des politiques d'activations sur les familles monoparentales pauvres. Il s'agit toutefois à ma connaissance de la seule expérimentation aléatoire en France de programme d'accompagnement de familles monoparentales. Ses résultats sont donc particulièrement pertinents dans les débats politiques récents, marqué par un tournant conservateur particulièrement violent à l'égard des plus vulnérables.

- **Codes Journal of Economic Literature** : I38, J16, J18
- **Mots clés**: Welfare-to-Work, single parents, active labour market policy, long-term unemployment

Welfare-to-what ? Experimental evaluation of an activation programme for single mothers in poverty in France

I want to thank Marc Gurgand for his supervision and guidance over the years and Karen Macours for her advices and support. I am grateful to Camille Terrier and Anne Boring whose comments in the pre-defense really helped improve this paper. This project and associated researchs would not have been possible without the initiatives and maintained efforts of Gabriel André, Bernard Tapie, Florence Thibault, Virgine Gimbert, Lucie Gonzales. I owe a great deal to Alex Galitzine and Nirsynne Nahhal whose stimulating discussions, corrections and proofreadings greatly improved this paper. Special thanks to Pedro Sant'Anna for sharing codes to adapt the did package to this setting. I am grateful to Bruno Palier, Antoine Bozio, Robin Huguenot-Noël, Michael Zemmour, Clement Carbonnier, Saad Loutfi and the participants of the Labour and public policies seminar of Paris School of Economics, the Séminaire Travail en Économie Politique (STEP) of Paris I Panthéon Sorbonne and the informal research group on poverty at LIEPP (Sciences Po) for their stimulating comments questions and suggestions at various steps of this project.

I Introduction

Over the past fifty years, the share of single-parent households have become increasingly prevalent in many OECD countries, now representing around one-fourth of all households in the European Union (Nieuwenhuis 2020). This trend is associated with increased risk factors for both parents and children, such as financial insecurity, isolation, lack of access to quality healthcare and childcare, or poor health and housing conditions (Broussard 2010; Duriancik and Goff 2019). Welfare States have implemented various policies but their level of generosity or conditionality remains at the heart of major tensions, involving antagonistic views on social justice, family dynamics, and the role of work in society, all through the lenses of gender, social origins, and racial origins (Moller 2002; McLanahan 2004; Reese 2005; Carcasson 2006; Foster 2008; Moffitt 2015; Reingold and Smith 2012; Mangin 2021; Herbst-Debby 2022).

Since the 1990s, OECD countries have shifted towards active labour market policies (ALMPs), with distinct waves of adoption characterised by neoliberal then inclusive growth narratives, broadly advocated by economists (Peterson 1997; Martin 2015; Crépon and van den Berg 2016; Peden 2017). Inclusive growth aligns with what Périvier and Sénac (2017) call a form of "*neoliberal morality*" through the *social investment paradigm* (Jenson 2010 ; Bonoli 2011 ; Morel, Palier, and Palme 2012). The latter try to reconcile economic efficiency with social justice concerns and shares with activation the idea that labour is the first protection against poverty. Social investments aim to both increase and use human capital to foster labour market participation through - among other things - education, training and job search assistance. However, it differs from activation by emphasising the idea that social inclusion cannot go without protection for the most vulnerable: the State must also play a role as a social buffer (Hemerijck 2014). Social investment had been an essential political framework in the European union over the past 20 years, leading Knijn, Martin, and Millar (2007) to argue that ALMPs have become the main policy framework for single parents.

While ALMP often have positive effects on average (Card, Kluve, and Weber 2018, 2010; Vooren et al. 2019), they have been criticised for their impact on vulnerable groups like single parents and their children (Ellwood 2000; Campbell et al. 2016; Avram, Brewer, and Salvatori 2018). For instance, the Cochrane systematic review of Gibson et al. (2018) summarises the effects of welfare-to-work programmes on single parents and reports almost no effect on employment, income or on mental health. Furthermore, and even when employment and income were higher for the lone parents in welfare-to-work, most participants continued to be poor. Despite large public spendings in activation policies, they often fail to alleviate joblessness and poverty among single-parent households, extending the risks to their children growing and remaining in poverty (Vandenbroucke and Vleminckx 2011 ; Løken, Lommerud, and Holm Reiso 2018; Rodríguez 2023).

In this paper, I evaluate the effects of an intensive welfare-to-work programme targeting single parents on long-term welfare in France. This programme called "*Reliance*" has been rolled out each year from 2018 to 2022 in the urban area of Nancy, the 16th largest urban area located in the North-East of France[1]. It targets single parents under 50 years-old registered for the French minimum income scheme: the *Revenu de solidarité active* (RSA). The initial stages of the support process involved diagnosing the main problems (such as over-indebtedness, housing, healthcare, and children's education), ensuring take-up of all social transfers and registration to the Employment agency. Throughout the year the intervention was organised alternately between group and individual sessions, at convenient times with regard to schools and daycare timetables. Participants involvement required about 15h/per week and, importantly, they could attend with their child(ren), for whom shared childcare duties among participants were organised in a dedicated space. The experimental sample is composed of 1671 single parents, 95 % women, 97% households living in poverty.

The *Reliance* programme is deeply rooted in the social investment paradigm by its target, policy objectives and design. It draws significant resources from various institutions[2], *investing* approximately €2,800 per participant - four times the usual spending per RSA recipient - in order to foster labour market participation and, ultimately, reduce poverty of highly vulnerable families. In the literature, these ingredients are usually predictive of higher effects on labour market participations. For instance, Bloom, Hill, and Riccio (2003) use data from three large-scale, multi-site random assignment experiments in the US and show that emphasis on quick return to employment,

[1] It was initially designed for three cohorts but two more were added because of the Covid-19 pandemic. See section II for details.

[2] Including the National family allowance funds and its local branch, the Caisse des dépôts, and the Departmental council.

personalised support, limited use of basic education and low staff caseloads increase effect size on employment and incomes. However the systematic review of Gorey (2009) shows positive effects are only found when the programme involve access to affordable childcare and decline when unemployment rates rise and jobs become harder to find. Conversely, the review of the effects of ALMP for women in Europe by Bergemann and Van Den Berg (2008) reports mostly positive effects, stronger for women than men, especially for training programmes. However, the vast majority of these research (39) are quasi-experiments while the only 4 randomised trials give opposite results (two positives, two negatives).

Relying on this state-of-the-art welfare-to-work programme for single parents in poverty, the main challenge is to ensure an evaluation that truly inform public policies. The analysis is thus guided by the following causal questions:

Does the programme increase labour market participation ? Does it reduce poverty ?

To answer that, I designed a staggered randomised experiment to assign encouragement each year using a block-random assignment defined by the product set of the number of children, unemployment registration and number of years on welfare. I focus on the four first cohorts of approximately 417.75 households each, with a total encouragement group of 843 households, and 828 controls. The former were invited to participate via formal letters, SMS reminders and phone calls. Throughout the years, the recruiting process has been adapted to foster participation. Letters were made more threatening, meetings were organised in the programme's meeting room, changed from collective information sessions to face-to-face individual meetings with project managers and former participants providing testimonies. On average, the take-up rate was about 38%, increasing from 28% to 47% between the first and fourth cohort. Participation is higher among the poorest, the least educated and those already registered at the Employment agency. I use matched administrative data from the National family allowance fund from January, 2017 to June, 2023, creating a panel dataset of 102749 observations to measure the effect of the programme on employment and poverty. I support and discuss the quantitative analysis with results from a qualitative evaluation conducted by an independent team of consultants (FORS 2020).

The results of this experiment are highly relevant to the French political debate, the past and forthcoming reforms. During its implementation, this programme was deemed very promising. It is praised in several official reports including those of the evaluation of the "Anti-poverty Strategy" launched by president Macron in 2018. It received the visit of the Minister of Health and Social affairs and a Secretary of State and quoted in MPs reports. However, social investment have been dropped from the government's narrative which, in a few years, took an unprecedented conservative turn[3]. After several waves of reforms tightening unemployment insurance eligibility and sanction measures, the French government[4] plans to enforce workfare obligations for all RSA recipients and already adopted several reforms towards it. As of January 2024, the Employment agency relabelled *France travail* is to oversee all RSA recipients by 2025, mandating 15 hours of work or social support under risks of monetary sanctions.

Contrary to this experiment, this new policy is more coercive and far cheaper. Yet this programme's effects can serve as an upper bound for the expected effects of the reform on poor single mothers. Indeed, There is a dearth of quality evidence on the effect of welfare-to-work programmes in France (Bono et al. 2021). In a review of the effect of social support on various dimensions, Cervera et al. (2017) note just a handful quantitative analyses in France. The limited exceptions mainly focus on employment outcomes for the unemployed or young adults. A systematic review of the literature on the effect of ALMPs on long-term unemployed individuals between 2000 and 2015 by Abadia et al. (2017) reports only one study involving welfare-recipients of the former minimum income scheme (RMI), before the introduction of in-work benefits (Crepon et al. 2013).

This study presents a unique experimental evaluation and contributes to filling this hole in the literature on welfare-to-work policies for single parents in poverty. While it is always a researcher's duty to provide clear and high quality evidence with transparent empirical method, the stakes in this study feel notably higher. To uphold the research's integrity, all analyses and outcomes were preregistered, thereby mitigating potential political pressures and preventing cherry-picking. In addition, this article is knitted using Rmarkdown, and all codes generating results are embedded in the files to ensure replicability and transparency.

By design, random assignment identifies the intention-to-treat effects of the programme and the effects on participants when used as an instrument for enrollment and excluding other causal path between encouragement

[3] For discussions on these shifts around the 2022 election, see for instance Knapp (2022),Hewlett and Kuhn (2022) or Durovic (2023)

[4] https://travail-emploi.gouv.fr/

and outcomes. The staggered entry of cohorts and the treatment effect dynamics by time-to-event require careful estimation strategy. Staggered designs have been the focus of many recent methodological contributions[5]. I use stacked regressions with inverse propensity score weighting to estimate the intention-to-treat parameters, interacting treatment effects with relative months dummies and a full set of block $\times$ months fixed effects. These interactions saturate the model and ensure comparisons with clean controls (Sun and Abraham 2020; Freedman et al. 2023). As a robustness check and mean to compare experimental estimates with second best identification strategies, I adapt results of Callaway and Sant'Anna (2020a) to aggregate month-cohort specific treatment effects to this experimental setting. As for instrumental variables, I use results of Borusyak, Hull, and Jaravel (2022) on shift share IV and simply demean the instrument with the block-specific share of assignment and retrieve the treatment effect on the treated due to one-sided non-compliance. I use cluster-robust standard errors at the block level and control the family-wise-error rate using either the Holm–Bonferroni correction or wild-cluster bootstrap (Hothorn, Bretz, and Westfall 2008; Alberto Abadie et al. 2022; C. de Chaisemartin and Ramirez-Cuellar 2022; MacKinnon, Nielsen, and Webb 2023).

In brief, I find a strong lock-in effect for the first half of the programme that reduces employment on the treated by -10 percentage points, income per capita by about € 85 each month and slows the climb out of poverty. These effects dissipates by the end of the training and I find no average effect on employment and incomes for at least a year after the end of training for the four first cohorts. Over that period, only 10% of the sample earned more than the poverty line and the distribution of potential income per capita of treated and untreated compliers are very similar.

Patterns of treatment effect are heterogeneous across cohorts until the end of the programme, but this heterogeneity stems mostly from differences in the counterfactual. Covid-19 and the increase of in-work benefits may explain these short-lived differences. However, at the end of the programme, the effects on employment and incomes are homogeneous across cohorts. In addition, the treatment effects on incomes and employment for those who had one child and those who had three or more children are reversed and puzzling: For those with one child, the programme reduces employment but has no effect on incomes ; For those who had three or more children, the programme seems to increase employment but reduce incomes, although both are imprecisely estimated. These results suggests heterogeneous changes in the composition of disposable incomes between families of one and three children. The programme gradually increases total cash transfers among participants. However, this effect is mostly mediated by changes in household structures.

As-treated analyses show large selection effects and the next best identification strategy using matched difference-in-differences with the control group rejects the experimental estimates from the 95% confidence interval. I conclude that the programme attracted those with the highest employment potential but slowed their re-entry into employment (lock-in effects), inducing significantly lower disposable income until the end of the programme, and no effect after. Although the programme selected those with highest employment potential, labour market participation is always lower than 40 % and very few exit poverty.

These results are consistent with a large strand of literature showing that active labour market policies, especially workfare and welfare-to-work programmes may worsen the situation of vulnerable single parents (Smedslund 2006; Gorey 2009; Mogstad and Pronzato 2012; Brady and Cook 2015; Campbell et al. 2016; Gibson et al. 2018; Avram, Brewer, and Salvatori 2018; Johnsen and Reiso 2020). Economics as a profession has a lot of responsibilities in the way society thinks of its problems. As Hirschman and Berman (2014) notes, "*the spread of economic discourse reshapes how non-economist policymakers understand a given issue. The spread of economists' technical tools determines the information available to policymakers and changes the process of decision-making*". Economists' narrative on welfare-to-work have been highly consensual despite many negative results. With more and more robust evidence coming-up with opposite effects from what policy makers intended, or economists thought would happen, perhaps the time is right for a paradigm shift.

The remainder of this research is organised as follows. Section II introduces the institutional setting and the intervention. Section III describes the experimental design, data and descriptive analysis. Section IV discusses the identification and estimation strategies. Section V presents the effects of the programme on the main outcomes and Section VI explores plausible mechanisms through a heterogeneity analysis. The interpretation, limits and policy implications of these results are discussed in section VII.

[5] See for instance the recent reviews by Roth et al. (2023).

II Supporting single mothers in poverty in France: Context and intervention

The programme targets single parents on long-term welfare in France through intensive social support. To understand the main difference between the treatment and the counterfactual, this section presents the general framework of active labour market policies for single mothers in France and discuss how this population fare under this regime. Then, I present the programme in detail and contrast it with the counterfactual and a workfare reform expected in 2025.

II.1 Single parents and the French minimum income scheme (RSA)

Before 2008, ALMP primarily aimed at reducing labour costs, particularly for low wages, and increasing employment rates through the sharing of working hours. The 35-hour work-week and the introduction of part-time and short-term contracts, as per Kramarz, Nimier-David, and Delemotte (2022), notably improved women's labour market participation, especially for married women. The pivotal turn in 2008 introduced a new welfare regime with in-work benefits and mandatory job search or social support, known as the Active Solidarity Income (*Revenu de solidarité active*).

The French *active* minimum income scheme (RSA): RSA merges the former minimum income scheme and the single-parent allowance, replacing the latter with the *RSA Majoré* for isolated parents for up to 1.5 years or until the youngest child is older than 3. Importantly, it includes two activation policies:

- "*RSA activité*," relabeled "*prime d'activité*" in 2016 (PA), is an in-work benefit inspired by the US Earned Income Tax Credit. It provides financial incentives for workforce re-entry through a monetary transfer, depending on household earned incomes, size, composition, and other social transfers.
- "Mandatory" social support and/or monitored active job search, managed by Departmental councils.

In 2020, around 2 million households received the RSA (Revenu de Solidarité Active) with a total cost of 12.6 billion euros (DREES 2022), roughly the same budget allocated to research, higher education, and innovation (13.4 billion euros in that year[6]).

RSA and PA are differential transfers with baseline levels assigned to eligible households, adjusted based on household structures and changes in earned incomes every quarter. Inspired by the Flat Tax model and the US earned income tax credit, the RSA design aimed to address the low incentives to work of the previous welfare system (Gurgand and Margolis 2008).

Labour market institutions underwent significant reforms in the Hollande and Macron presidencies, notably reducing both employment and unemployment protections (Milner 2017; Leruth 2017; Gazier 2019). In 2019, the French government increased monetary incentives following the Yellow Vest movement, widening income eligibility thresholds, increasing baseline amounts, and offering higher individual bonuses in the PA reform. This reform, estimated by Dardier, Doan, and Lhermet (2022), led to a 37% take-up increase and an average gain of € 70 per month.

Despite France's reputation for a generous Welfare State, 66% of RSA household live under the poverty threshold $i.e.$ less than € 1 140 per consumption unit[7], more than four times the general population's average of 14.8%. This population is highly isolated, with 49% not seeing friends or family in the previous month and 26% having no contact at all. Health problems are prevalent, with 21% reporting poor or very poor health, 43% having at least one chronic illness, and 38% stating it limits their capacity. Additionally, 22% are at risk of depression, and 15% had to forego seeing a doctor in the previous year due to economic reasons (DREES 2022, chap. 15 and 16). These statistics reflect the most recent reality, with documented deterioration since the COVID-19 crisis according to CNLE (*Conseil National de Lutte contre l'Exclusion*) (Duvoux and Lelièvre 2021).

6 See the Senate report

7 In 2019, the monetary poverty threshold computed as 60% of the median standard of living. See https://www.insee.fr/fr/statistiques/7710966.

Single mothers in France Single-parent families constitute 90% of beneficiaries of the "RSA majoré" (increased RSA), with a staggering 97% being women, and they make up 43% of RSA beneficiaries, 89% of whom are women. Female RSA recipients, particularly single women with children, face a monetary poverty rate of 73%, significantly higher than their male counterparts. The prolonged duration in RSA is associated with a decline in their standard of living, a trend worsening after the third year (Cour des comptes 2022, chap. 3).

Health issues, both physical and mental, often become a significant barrier to sustained employment or re-entering the workforce. Research by Delattre, Moussa, and Sabatier (2019), using French panel data, demonstrates a strong reciprocal effect of health status and employment, revealing a cycle where health deterioration at time t-1 predicts employment at time t, and past employment status influences health. RSA is sometimes used as a buffer allowance before accessing disability allowance, and the intersection of poverty, welfare eligibility, and disability is evident, with a quarter of new disability allowance recipients having previously received RSA benefits. However, recognizing their disability is challenging due to a lengthy process and unclear eligibility criteria, trapping them in a situation that hinders compliance with welfare-to-work obligations, leaving them in an uncertain status between the two institutions (Cattoen et al. 2022).

Moreover, the life trajectories of single mothers are often marked by traumatic experiences, including exposure to violence, domestic violence, and sexual assaults, extending back to their childhood. Studies, such as the analysis of data from the VIRAGE survey by Brown (2020) in France, reveal a high prevalence of violence, with its effects persisting post-separation. Notably, women no longer in relationships report experiencing harm more frequently, with a significant proportion enduring severe violence, often rejecting conventional marital norms and embracing enduring single parenthood.

Despite many being employed or having worked in the past, victims of violence often face higher instances of unemployment or inactivity for periods exceeding six months. Sometimes, Intimate partners use violence and coercion (IPVC) to obstruct women's economic independence[8]. Such violence contributes to increased vulnerability, self-deprecation, and even depression, leading to pronounced social isolation. Consequently, these women often feel illegitimate in returning to employment, especially with limited work experience.

II.2 Intensive social support for single parents on long-term welfare

The **Reliance** programme is an experimental comprehensive support initiative to foster "*sustainable reintegration into employment and society*" of single parents on long-term welfare[9]. It was funded and implemented by the Departmental council, the *Caisse des allocations familiales* (CAF) of Meurthe-et-Moselle, the National family allowance fund (Cnaf), and the Caisse des Dépôts. The programme targets eligible population residing in the urban area known as *Grand Nancy*[10], which encompasses 20 municipalities[11] and includes 8 neighbourhoods categorised as deprived for a total population of 420,120 inhabitants in 2019, of which 47,799 lived in single-parent families (INSEE 2023).

[8] For instance, Riger and Krieglstein (2000) report that, within a job programme in the USA, 47% of women who were victims of violence mentioned their intimate partners attempting to prevent them from pursuing education or training. Both victims and non-victims in this sample were discouraged from working by their partners. Spencer et al. (2020) analyse the effect of TANF for victims of IPV and also find such effects. For a review of the links between violence and family type, see Tur-Prats (2019).

[9] A web-page presenting the programme is still active in February 2024 : https://www.arelia-asso.fr/index.php/25-canevas/newsletterlirelasuite/510-reliance-est-un-nouveau-dispositif-experimental

[10] It is located in the heart of the Grand-Est region, which had 1,918,000 salaried jobs by the end of 2020. The employment dynamics of the region are mixed, with a net loss of 50,000 jobs from 2010 to 2020, primarily driven by declines in the industrial (-13.9%) and construction sectors (-8.1%). The tertiary sector remained relatively stable.

[11] Art-sur-Meurthe, Dommartemont, Essey-lès-Nancy, Fléville-devant-Nancy, Heillecourt, Houdemont, Jarville-la-Malgrange, Laneuveville-devant-Nancy, Laxou, Ludres, Malzéville, Maxéville, Nancy, Pulnoy, Saint-Max, Saulxures-lès-Nancy, Seichamps, Tomblaine, Vandoeuvre-lès-Nancy, Villers-lès-Nancy.

A welfare-to-work programme in the social investment paradigm The programme takes places in a renovated building in Vandœuvre-Les-Nancy, on the main tram-line at the heart of a neighbourhood prioritised for urban policy, equipped with offices for interviews, meeting rooms for group sessions, a children area, a communal kitchen, and computers. It is a delegation of a public service mission to private but non-profit operators: three well-established local associations. This experiment thus shares some common features with researches that compare public and private provision of workfares (e.g. Behaghel, Crépon, and Gurgand 2014). Social support is implemented by three experienced[12] social workers and an executive, for 82 participants per year, on average. This caseload per social worker is thus very low; for comparison, Jacquey-Vazquez (2017) reported that in 2015, The Employment agency intensive support advisors were each responsible for an average of 108 unemployed individuals. Participants enroll in the programme for a year hoping to find adapted and sustainable solutions to employment and precariousness issues, making the concept of social investment central to this experiment. The assumptions regarding the effects of the programme are based on both a *capacity-building* approach – where participants benefit from the programme by developing or maintaining skills, building a network, etc. – and an *emancipatory* approach that seeks to alleviate the specific burdens and obstacles faced by each family.

The support is based on a comprehensive assessment of beneficiaries' situations (personal, family), professional backgrounds, training needs, and environment. The initial stages of support involve diagnosing the main issues the family faces (overindebtedness, housing, health, children's education, etc.) and ensuring access to rights by resorting to national and local aids, including visits of social workers from the Family allowance fund. This step also aims at overcoming peripheral barriers to employment such as facilitating access to childcare for the families involved and confidence in their level of social transfers if they take a job. One objective is to quickly alleviate individuals' mental load, enabling them to regain their own resources (Mani et al. 2013; Schilbach, Schofield, and Mullainathan 2016).

In this project, evaluation is a central component and has been conceived as part of the programme itself. Beyond my research, the Departmental council funded a qualitative analysis (FORS 2020) and coordinated two masters dissertations: one in sociology on how social workers perceive innovations in their work (Mahdi 2021), and one on non-take-up of social programmes (Chachou 2019). We also conducted surveys in order to measure the effects of the programme on more subjective dimensions[13].

The social investment perspective legitimises public intervention through a demonstration of its social returns in the medium to long term (Périvier and Sénac 2017). The programme would be considered a success if it generated positive "social returns" through "avoided costs" in social benefits. In this regard, the average cost per participant is estimated to be around € 2800, approximately four times the average expenditure for regular support (Mahdi 2021).

A highly promoted innovation Recent official reports have been highly critical of the implementation of job search monitoring and mandatory social support for RSA recipients (Pitollat and Klein 2018, Aout; Damon 2018; Cour des comptes 2022). Concerns include inadequate support reaching eligible recipients, a time-consuming referral and support process, heterogeneous content, and insufficient employment outcomes. This experiment takes a totally different path with intensive support, innovative recruiting process and a strong emphasis on measuring its effects. Specifics dully noted in the aforementioned reports, who cite this experiment as a promising way forward.

It also received significant political and media support, including an official visit[14] by the Minister of Solidarity and Health, Agnès Buzyn, and the associated Secretary of State, Christelle Dubos. The programme was awarded the Afigèse[15] 2021 Prize in the "public policy evaluation" category, despite the fact that the results of the impact evaluation had not yet been released. It has been featured in several articles in regional newspapers, radio broadcasts, and other specialised media outlets.

[12] They all have a master degree and several years of work experience.
[13] The objective is to measure relatively long-term outcomes and taking into account funding constraints, we decided to survey the cohorts at the same date and to repeat the questioning respectively 2 and 3 times for the 2019 and 2020 cohorts to see if these results evolve. The analysis of the responses to these surveys will be the subject of another publication.
[14] Meurthe-et-moselle.fr Actu - visite ministre
[15] *Association finances, gestion et évaluation des collectivités territoriales*, a network that brings together professionals in local public finance and management, public policy evaluation, and territorial public management. Press release

A highly vulnerable target population The target population is highly vulnerable. At the time of random assignment, the average standard of living is about 718 €$_{2015}$ and only 2.8 % are above the poverty line in 2019. In the qualitative evaluation (see details in section III), most single parents report health problems, employability constraints, budgetary difficulties, and housing instability. Often less educated and engaged in precarious employment, particularly those with migration backgrounds, they often withdrew from the labour market after the first child's arrival, entering RSA. They now juggle domestic responsibilities and parenting alone.

Facing economic hardship and social isolation, compounded by the dual stigma of being RSA recipients and single parents, they often experience a sense of "social shame," exacerbated in cases of prior violence. The qualitative evaluation underscores their vulnerability to life's setbacks, lacking social and economic safety nets, and being more exposed to health and family-related risks. Some had their children removed from custody or are under monitoring from children protection services. Several experienced intimate partner violence and coercion, which, for some, led to depression and post-traumatic stress disorder of the victims and sometimes their children. In this sample, only 20% receive child support from the other parent of their children.

Recognising the unique challenges of this population, Reliance provides comprehensive support, combining individual and group assistance. Thematic workshops cover diverse topics, involving local community partners, field trips, and leisurely outings. Workshops address benefit rights[16], parenthood, and digital literacy, aiming to humanise administrative procedures and alleviate emotional burdens. They offer practical assistance with applications for social housing, affordable school lunch, and suitable childcare options, as well as strategies for daily life organisation. Additionally, workshops explore self-awareness, relationships, and gender norms.

Complementary workshops cater to individuals' interests and backgrounds, including activities related to specific fields, outings, and well-being. Support days alternate between group and individual sessions, accommodating participants with or without children. Children have a designated area and are supervised by participating parents. The schedules align with school hours and participant availability, evolving based on observations and needs. Activities primarily focus on creating and validating realistic professional projects, addressing steps like education, internships, and improving job search efficiency, aligning expectations with job opportunities.

A staggered roll-out marked by the Covid-19 crisis To take the social investment perspective seriously, estimates of the causal effects of the programme are required. At the onset, we wanted to recruit a 100 participants every year, and policymakers defined the *success* of the programme as a 10 percentage point increase in employment. I conducted simple power computations with an *optimistic* take-up[17] of .5, 12 % attrition, mean employment change in control group of 10p p (SD=.3), 80 % power and 5 % two-sided test, reaching approximately 1200 households. We started with a target of three cohorts of 400 households.

Each cohort was randomly sampled and assigned to treatment. In general, caseworkers refer welfare recipients to such programmes, making comparisons with untreated recipients unlikely to recover causal effects of interest. This recruitment process is thus very different from common practices. In fact, it changes the pool of participants and reaches welfare recipients who would not have come otherwise. The qualitative evaluation strongly emphasise this results and documents a swift swing of opinion regarding random assignment from previously sceptical social workers. Surveying participants and non participants, FORS (2020) also reports a strong support from participants themselves.

Through this period, the economic environment was affected by the COVID-19 pandemic, and the increased in-work benefits in the 2019 reform of PA. The pandemic disrupted the implementation of the programme for the 2020 cohort and the economy was almost entirely shut-down for a while, before slowly recovering[18]. As a consequences, we opted for the pursuit of the programme and secured funding for two additional cohorts. We enlarged the sample for the 2021 cohort to increase precision. In 2022, the pool of eligible families that had not been already sampled was too small, especially for long-term recipients. To build the 2022 cohort, we sampled the remaining eligible population and random samples from the control groups of the previous cohort. From November 2021, composition of the control groups of the 4 previous cohorts change. In the end, the intervention has been rolled-out from 2018 to 2022 in a staggered design summarised in Figure A.11 in the Appendix.

[16] This component of the programme turned-out to be very important as shown in Galitzine and Heim (2024).
[17] This parameter was provided by project managers who expected a large enthusiasm from participants.
[18] We give more details of the adaptation and consequences of the pandemic in Appendix A.II.

III Data, experimental design and descriptive statistics

III.1 Data sources and main variables of interest

Matched administrative records The design of this experiment is based on repeated draws of random samples of eligible households from administrative records of the Departmental council. These samples constitute cohorts for which treatment has been randomly assigned following the protocol described in subsection III.2. The initial datasets have been supplemented with these design variables and matched with monthly administrative records from the National family allowance fund (CNAF). The ALLSTAT files describe the situation of every beneficiaries for a given month. They contain information on the "*household heads*" and their possible spouse (including gender, year of birth, marital status, activity status, nationality, and so on), their dependent children (years of birth, alternated residence status for family benefits, absence of a parent, the legal benefits they receive (the individual social action aids they have benefited from), and several measures of household incomes. Details on the data sources and variables used in this analysis are presented in the Appendix B. Codes to generate these variables are available upon request. The main drawback is that they are not meant to measure employment but record relevant informations to compute all social transfers. However, they provide very good quality measures of incomes for the poorest and are consolidated over 6 months to account for treatment delayed, controls and adjustments.

I construct a baseline database of 2662 households from five cohorts measured the month before random assignment, including 548 households excluded from the experiment and 53 from the pilot study (see section III.2 below). The *experimental sample* consists of 2073 households. Additionally, panel data from the National family allowance fund and data quality assessments reveal complete information for 92% of the post-randomisation period, with only 4.1% files lost, showing no differential attrition between treatment arms.

Qualitative evaluation FORS, a consulting firm specialised in social sciences and public policy evaluation *to la française*[19], conducted the qualitative evaluation of the programme under the sponsorship of the Departemental council, with the report finalised by the end of 2020. Throughout my analysis, I incorporate insights[20] from this report to support and contextualise my findings, referencing it as the qualitative evaluation or citing it as FORS (2020). The methodology was organised as follow: Initially, the mission was framed through meetings with programme pilots and the Reliance team, alongside documentary analysis. Then, the framework and survey tools were validated to ensure alignment with research objectives. The investigation unfolded in two main phases: the first involved interviews with participants from Cohorts 2018 and 2019, as well as focus groups and team meetings. The second phase expanded the scope to include follow-up interviews with Cohort 2019 participants, interviews with Cohort 2020 participants, and discussions with individuals who declined programme participation, along with interviews with partner organisations. The final phase encompassed the synthesis of findings and recommendations, including additional interviews with Cohort 3 participants and a collective interview with the Reliance team. Adjustments were made due to the COVID-19 pandemic, necessitating alternative data collection methods such as video conferences and telephone interviews, while still adhering to the research objectives and participant preferences.

Main variables of interest Households are identified with a unique id[21] and the cohort to which they belong. The Department's database contains the blocks' identification and associated variables[22], the encouragement (Z) and treatment (D) status, the date of randomisation from which we construct time-to-event variables. I also keep the Department's social workers' assessments (favourable or reserved).

[19] See Delahais and Lacouette-Fougère (2019)

[20] Translations of direct quotes from interviewees are my own.

[21] Cnaf does not have a national id for each household, identification is specific to the CAF families that are registered. When they move to another county or their relationship status changes, the household ID changes. I retrieved households that moved using other identification variables in the main information system.

[22] Number of children (1, 2, 3+), registration to the Employment agency (True/False), and length of time receiving RSA (2 to 5, 5 to 10, 10 or more years) ;

The main outcomes have been defined in the registration plan[23]. Table B.4 in the Appendix describes how their were constructed and alternative measurements used as robustness checks. This paper focus on two main outcomes in line with the main research questions:

- **Employment**, measured by reported labour incomes
- **Poverty**, measured by disposable incomes per consumption units and compared with the poverty threshold.

In the pre-registration, I also registered welfare and in-work benefit eligibility and total social transfers. Their analyses are presented in the appendix together with alternative measurement of the main variables, and briefly discussed together with the main results. In Galitzine and Heim (2024), I also investigate the effects of the programme on family structure (number of children and relationship status) and source of incomes.

Measures of labour market participation The main outcome of interest is labour market participation although I can only observe self-reported monthly household incomes used to compute social transfers. PA and RSA depends on all labour incomes in the households and other sources of incomes (including other social transfers) in the previous quarter. Each household must report their incomes by member of the household and type of income. I use these quarterly income reports and define employment as a dummy that equals one when the parent declares *her* positive labour income in a given month and zero otherwise. The main limitation is that I can only observe employment for families who report their quarterly incomes. Employment levels may be underestimated if those who stopped reporting their income did so because they earned more than the eligibility thresholds. As discussed in the previous section, these variables suffer from minor attrition which does not differ between the encouraged and control groups. Therefore, employment may be measured with error, but should not affect treatment effect estimations.

The database also contains detailed information on parents' "main" occupation, which can be classified according to the Labour force survey definition in 7 categories[24] and is also observed for parents who did not report their quarterly incomes. It has less missing cases but is less reliable nonetheless. Indeed, this variable is filled by CAF agents and there may be important delays, depending on controls or updates for motives unrelated to employment. Participants met with CAF social workers several times at the beginning of the programme and their files were likely updated while the control group was not. Those who do not report their quarterly incomes are even less likely to be updated on the right time. I only report estimates based on this variable as a robustness test in the Appendix.

Estimating the effect of the programme on poverty The Family allowance fund computes income per consumption units, which encompasses every source of income (labour and capital income, unemployment insurance, health pensions, and cash transfers from CAF) from every member of the household weighted by the number of consumption units. The first adult is weighted 1, any other member of the household that is above 14 years-old is weighted .5, and children under that age are weighted .3. Single-parent households receive .2 additional points. Note that this scale tend to overestimate the living standard of single parent households for whom economies of scales do not existent *a priori*. A 14 year old child is weighted the same way a partner is (Le Pape and Helfter 2023). In this research, I compare this variable with the poverty threshold from 2019, which is € 1140 according to INSEE. I convert incomes and threshold to 2015 values and estimate the effect of the programme on the share of households with less than this threshold to measure the effect on poverty.

[23] Note that in the registration plan, there are additional outcomes based on surveys which have not been exploited yet. My conclusions are based on this first set of hypotheses and the others will be analysed in another research, also involving data from the 2022 cohort and longer observation windows for the early cohorts.

[24] Other Inactive, Other Unemployed, Retired, Student Or Training, Unemployed, Unknown, Works

Covariates used in the analysis Blocking variables already accounts for the cross effects of number of children, registration at the Employment agency and years receiving RSA across cohorts. A large set of attributes is observable, further increased by panel data and the possibility to use past outcomes and characteristics as covariates. However, I remain parsimonious and only use a selected set of observables that correspond to the typical variables social workers would observe. The definition of these variables is presented in table B.5 in the Appendix. Following S. Athey and Imbens (2017b), I only use *dummified* versions of the covariates to keep a clear interpretation of all coefficients and limit bias from parametric restrictions.

Baseline outcomes typically predict subsequent outcomes and I include them measured the month before random assignment and centred. I use typical socio-demographic variables including dummies for French citizenship, quartiles of age, having children under 2, having children between 3 and 5 and having children older than 16. French citizen and younger worker usually face less constraints on the labour market (Hargreaves 2015; Anne et al. 2019), while children under three require intense care, preschool is mandatory from age three which may ease parents re-entry in the labour market (Goux and Maurin 2010). Older children may help care for younger and be more autonomous in general, thus possibly reducing parental constraints. Data from the Departmental council also include a measure of education that is known for approximately 80.5% of the experimental sample. I construct a simple High/low/unknown education variable, considering high school diploma and above as high education. To control for incomes, I use a dummy for taxable incomes 2 years before higher than the median[25], quartiles of disposable income per consumption unit at baseline, and dummies for housing benefits, public alimony (ASF) and receiving child support[26]. Finally, I define dummies for being re-sampled in the 2022 cohort and being assigned to the encouragement group, and allow specific trends for these groups.

III.2 Selection of participants

The design uses simple random sampling from administrative data among the eligible population on the one hand, and block-random assignment of encouragement on the other. This section discusses the recruitment process.

A) Sampling, pre-selection and random assignment

Each year, the Departmental council draws from its administrative databases a random sample of about 500 households that meet the following inclusion criteria:

- Single-parents under 50 receiving RSA and with at least one child present in the housing ;
- Resident of one of the municipalities of the *Métropole du grand Nancy*;
- Last RSA registration older than two years;

We ran a pilot study to test the programme's attractiveness and randomly selected 53 households in November 2017 and invited 41 of them to participate in the programme. Among them, 15 enrolled, giving us an idea of the expected take-up ($\approx$ 36.6 %). These families started the programme in January 2018 for a few months in a very small group before being joined by the participants of the 2018 cohort in April 2018[27].

[25] Which means no taxable incomes.

[26] I do not use early childhood allowance which is colinear with children under 3 and Family allowance and Family supplement which are colinear with number of children. Indeed, Family benefits are open from the second child, and the Family supplement from 3 children.

[27] Figure B.14 in the appendix shows the share of single parents with positive labour income in the pilot and the 2018 average. This first very small group had spectacular outcomes showing up very early. However, they lost their jobs during the pandemic and did not catch up since.

Caseworkers' orientation: A preselection After extracting the 500 files, civil servants of the job-support bureau of the Departmental council assessed files' *appropriateness* for the programme. They contacted each household's caseworkers and/or consulted their files to sort into three categories.

- "Favourable" are parents who would typically be referred to by their caseworkers for the programme had we used a more conventional recruiting approach.
- "Reserved" are generally formulated for families who have little contact with their caseworkers or if the known problems of the family make participation in such a programme *deemed unlikely* (typically because of known social or health problems)
- "Unfavourable" are given to families who already have a job or are already enrolled in another programme, to families followed by the child welfare service or to those for whom the programme is deemed inappropriate taking into account the information known to the services

The latter group represents an average of 21 % of the sample and were excluded from the experiment. The experimental sample is composed of households with favourable and reserved assessment. This pre-selection has an important consequence: the experimental sample is no longer representative of the population it was drawn from. At the same time, it is closer to ecological conditions and enables comparisons between those who would have been referred to the programme and those with *reserved* initial assessment, and how the treatment effect varies between these groups.

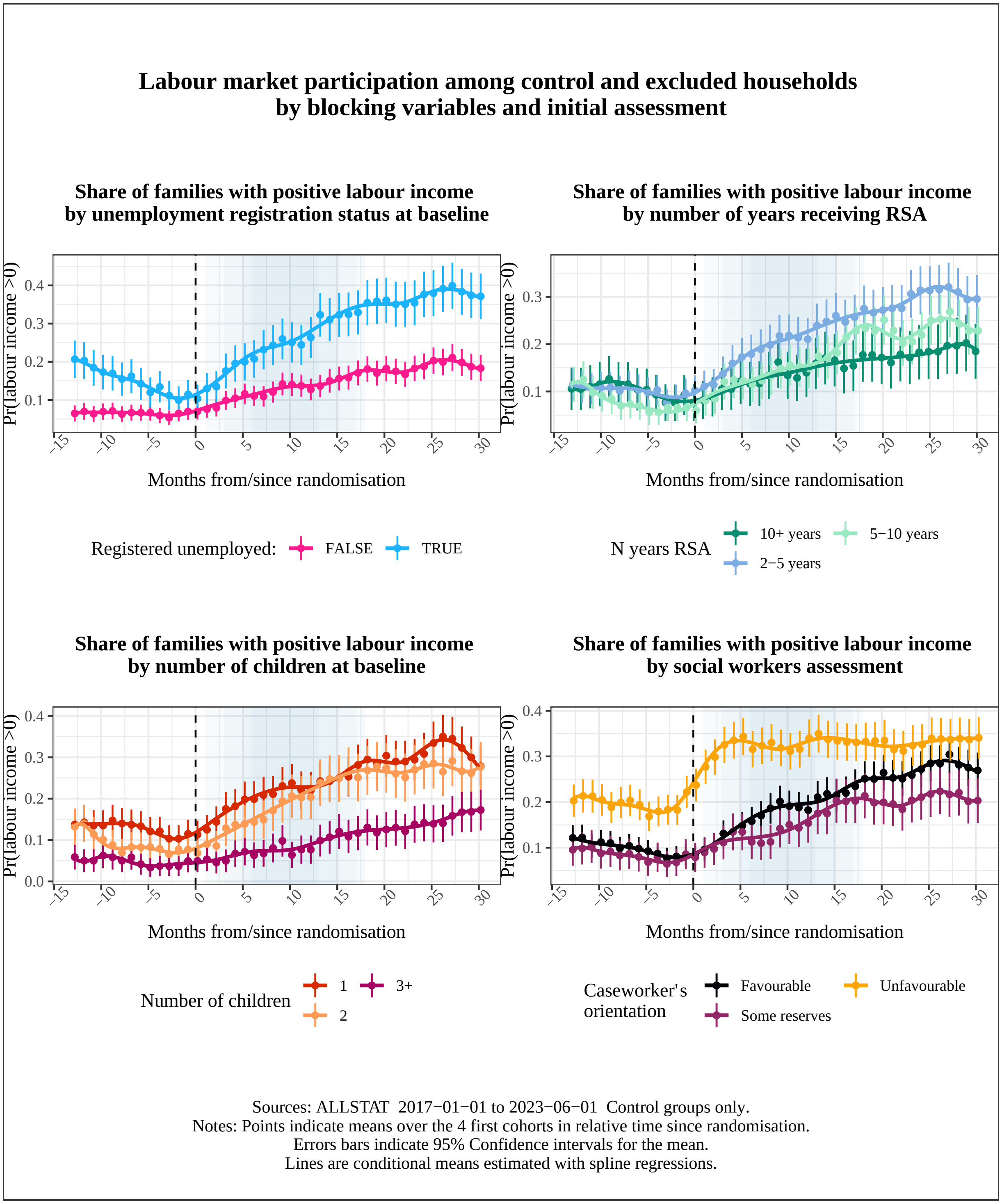
Randomisation protocol Each year, the Departmental council provided the random sample of eligible households with social workers' assessment and blocking variables. To assign encouragement, I use block-randomisation[28] within strata based on the cross product of:

- Registration status to the Employment agency (True/False)
- Number of children at home (1, 2, 3 or more)
- Years receiving RSA (2 to 5, 5 to 10, more than 10 years)

The product set of these variables define single parents' *type* between which I expect different reaction to the encouragement, different outcomes and possibly heterogeneous treatment effects. Registered unemployed are expected to be closer to the labour market, the number of children increases constraints due to parental obligation while it has been shown that the longer people receive RSA, the less likely they are to find a job (Cour des comptes 2022). These hypotheses are confirmed in the data, as shown in Figure 1. Each panel illustrate the share of families in the control group with positive labour income based on the values of the blocking variables. It also shows the gap by social worker assessments including the excluded ones. The data are presented over relative months from randomisation, displaying mean values with 95% confidence intervals and smoothed conditional expectations estimated through splines. The key insight drawn from this visual analysis is that prior to random assignment, most groups exhibit comparable outcomes, and as time progress, the proportion of families with positive incomes increases but evolves distinctly across the different groups. Those registered at the Employment agency are significantly more likely to work before they are sampled and have a higher job finding rate and employment level than those unregistered. Files that were excluded from the experimental group are more likely to work prior to random assignment, and increase by more than 15pp in 6 months before reaching a plateau at about 1/3. Those with favourable assessments have a slightly higher job finding rate than those with reserved assessments.

[28] I draw and sort lotteries from a uniform random variable using Stata and treat the top first half, rounded above. Blocks' sizes vary within and between cohorts and the share of encouraged families is not exactly .5 in every stratum.

Figure 1: Blocking variables and caseworkers initial assessment predict heterogeneous employment trajectories



Balance check and descriptive statistics Table B.6 provides the means, standard deviations, and differences between the encouraged and control groups. As expected, the random assignment of encouragement successfully balanced the groups across all variables used in this analysis. It provides the average values of the main outcomes and covariates at the time of random assignment. The sample is composed of 95% women aged 36 on average, 80% are French citizens, living with €$_{2015}$ 710 per month per consumption unit. The average household income is about

$€_{2015}$ 1400 while the average amount of social benefit is $€_{2015}$ 1300. About 1/3 has a child below 2 and 1/3 have a child older than 16. Half of them have a high school degree or higher while we don't know the education level of 1/3 of them. Only 1/5 receive child support from the other parent and 65% receive the family support allowance instead.

B) Increasing efforts to foster participation

At the beginning of the experiment, we were unsure whether we should foster participation through incentives and nudges or resort to more sever approaches, including threats of sanctions and benefit reductions. In the first year, the encouragement group received an official letter with information on the programme emphasising the potential benefits of participation and inviting them to public meetings where they could obtain more information. Following these meetings, interested families were invited to an individual appointment to sign a contract. Throughout the experiment, enrolment was voluntary, and parents could refuse to attend the meetings or enter the programme after the meeting. The take-up in the first cohort was the lowest[29]: 27.9 %.

We decided, in accordance with the international review board, to adapt the recruiting process for the subsequent cohorts. In particular, we opted for an ambiguous yet slightly more threatening letter as a way to foster attendance to the information meetings. From 2019 onwards, the invitation letters were sent earlier adding a bold sentence to the letter in French (see the template A.10 in Appendix A.I):

"I hereby inform you that **your participation in this meeting aligns with your obligations under the legislation of the RSA programme.**"

Our goal was to attract as many participants as possible to the informational meetings (by imbuing them with a perceived sense of obligation). However, the meeting place was moved from the Departmental council to the programme's meeting room, allowing prospective participants to visit the place and receive a warmer welcome[30]. The meetings retained their role of disseminating information and emphasising the benefits of the programme.

Moreover, former participants attended these sessions to share their experiences. Indeed, the content of the meetings also aimed to mitigate mistrust, diminish stigmas, and dispel any ambiguity regarding the mandatory nature of participation. Attendance at the meetings increased as a result, leading to a significant rise in the participation rate, reaching 36.6 %. At that time, the preliminary results of the qualitative evaluation revealed that the main motivations for participants was actually the fear of sanctions, and not potential benefits. In focus groups of participants, all agree that "*people would not come otherwise*", but stress the need for the programme to remain optional. An information only made clear during the meetings. Almost all individuals interviewed by the qualitative research team perceived their attendance as obligatory, or at least recognised a risk of their benefits being cut[31]. While the number of people who actually attended the collective meetings was sufficient, only 66 %[32] followed through with the programme later on. The project leaders estimated that an invitation to an individual interview organised by one of the social workers

[29] Note that compared with other comparable experiments in France such as Crepon et al. (2013) and Castell et al. (2022), this participation rate is rather high.

[30] A qualitative study was conducted in 2019 and 2020 surveying single parents of the three first cohorts (FORS 2020). It reports that these changes seemed to promote participation and strengthened the beneficiaries' commitment, both at the ontset and throughout the programme: the friendly environment, cleanliness and suitability of the premises, their aesthetics, etc. Participants emphasise material details that create a human and warm environment (such as coffee, for example), as well as more intangible aspects such as listening, kindness, and the "non-condescending" and non-judgmental attitude of social workers during these meetings. These elements form a framework that, according to participants, sets it apart from other social or support schemes they encountered in the past.

[31] For instance:

- "*I had received a letter. It was an information meeting, it was a requirement due to my RSA, I was obligated to attend, we had an attendance sheet, to say 'I came and I sign.' And afterward, it was our choice to stay or not. For me, I understood it that way.*" (Female, 44 years old, 2 children, on RSA since 2011)
- "*When I read the letter from CAF, I thought it was mandatory to do it, it was poorly presented. Otherwise, I thought I would lose all my RSA. It was poorly expressed, I think.*" (Female, 41 years old, 3 children, separated for 2 years, on RSA since 2009)
- "*We received a letter from the Departmental council, stating that it was a programme for single-parent families on RSA. It was not an obligation, but behind it, we knew there was a trap, your RSA could possibly be affected.*" (J., 39 years old, 3 children, on RSA for over 5 years).

[32] To comply with the General Data Protection Regulation, attendance sheets were only used by project managers to later contact participants who wanted to sign-in and destroyed after recruitment was complete.

of the project and former participants would be more likely to increase the participation rate by better explaining the interest in participating in this programme. Additionally, collective information sessions made some individuals uncomfortable, leading them to opt out. Individual interviews were expected to reduce judgement, comparison and stigma, a phenomenon often observed in the first waves, sometimes resulting in the feeling that the programme is not suitable for them. To address this issue, we changed the recruitment process from public sessions to face-to-face meetings with social workers from the programme and former participants. They provided testimonies and feedbacks of their own experience in the programme. This change in the recruiting process has been approved by the IRB and motivated by the fact that seeing the pool of other welfare recipients promoted stereotypical views and discouraged some from registering (FORS 2020).

Individuals who have experienced similar situations are perceived as more credible models than "institutional" voices in stimulating change among their peers: they have a better ability to establish empathetic relationships, convince others, and serve as role models. Mrs. D., a participant from the 2018 cohort and a "witness," explains: "*We realised that when it's presented by a former participant (me), it's perceived differently. It's amazing, this difference, the listening is different, we quickly noticed it.*" Furthermore, using the testimony of former beneficiaries also allows them to be valued and highlights their skills and experiences. For instance, a participant from the 2019 cohort explains: "*She explained her past to me, honestly, hats off. And you feel less alone. We're not alone in this case. And there's no judgement, we feel like we're all equal. I felt welcomed. It gave me a boost, that's what I needed in the end*" (F, 48 years old, 4 adult children, on RSA for 3 years, met in November 2019 during a mobilization day). Another says: "*It gave me a little more confidence to have someone who came like me, it reassures me*" (F, 34 years old, 4 children, met during the mobilisation day).

This change induced increased recruitment costs and duration, but ultimately led to a higher take-up rate: for the 2020 and 2021 cohorts, the take-up was 42.1 % and 47.2 % respectively. The 2022 cohorts received the same letter than the previous cohorts but did not receive text message reminders. The take-up rate was then 37.4 %. In the 2022 cohort, some participants did not received the reminders, probably explaining the lower take-up.

Table 1: Average effects of encouragement on participation by cohort

	sample					
	Full sample	*Cohort 2018*	*Cohort 2019*	*Cohort 2020*	*Cohort 2021*	*Cohort 2022*
Encouragement	0.386***	0.279***	0.363***	0.421***	0.472***	0.378***
	(0.018)	(0.027)	(0.027)	(0.039)	(0.045)	(0.036)
Num.Obs.	2073	395	397	386	493	402
R2	0.282	0.195	0.249	0.300	0.354	0.266
R2 Adj.	0.249	0.156	0.214	0.266	0.329	0.228
Std.Errors	by: strataXc	by: strataXc	by: strataXc	by: strataXc	by: strataXc	by: strataXc
FE: strataXc	X	X	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sources: Sample: Cohorts 2018 to 2022 at the time of randomisation.
OLS regressions of participation on encouragement, recentered by within-block propensity scores and blocks x cohort FE. Cluster robust standard errors adjusted by blocks x cohorts in parenthesis.

Table 1 summarises the average effects of encouragement on participation and separates estimations by cohort[33].

[33] Simply regressing participation on encouragement and block fixed effects

The take-up dramatically increased over the years and there are 19.25 pp difference between the first and fourth cohort. This substantially increases power and precision. However, it has the potential to alter participant profiles by mobilising those more averse to sanctions or by modifying motivations and thus individuals' actions (Redman 2020). This is also why testimonials and reinforcement of information about the optional nature of the programme were concurrently implemented. In any case, we managed to involve almost 4/10 single mothers in the encouragement groups.

III.3 Profiling the population of compliers

The average effect of encouragement on participation over the 4 cohorts is about 38.6 pp and we already saw that there is heterogeneity across cohorts. To understand the selection process, I start by estimating heterogeneity in the effect of encouragement on participation for a selected set of covariates: the same I use in all models. More precisely, I want to know if, *ceteris paribus*, the effect of encouragement is higher than the average effects for different groups of individual. To do that, I follow Lin (2013) and simply add de-meaned selected covariates and interactions with encouragement in the first stage. Denoting D for participation, Z for encouragement, I estimate the following regression using OLS (Lin 2013):

$$D_{ijc} = \alpha + \pi Z_{ijc} + \dot{\mathbf{X}}'\boldsymbol{\rho} + Z_{ijc}\dot{\mathbf{X}}'\boldsymbol{\beta} + \varepsilon_{ijc} \tag{1}$$

Where elements of $\dot{\mathbf{X}}$ are $\dot{X}_{ijc} = X_{ijc} - \bar{X}$ are the covariate deviations from the sample mean. Furthermore, I follow S. Athey and Imbens (2017a) and use only binary variables so this model is fully saturated and estimate the conditional expectation function perfectly. In this model, I don't use block fixed effect but add the blocking variable separately to estimate the average difference across variables. I weight observations by the inverse instrument-propensity score $\hat{q}_{jc}$ estimated[34] using a probit of encouragement on block x cohort fixed effects to ensure conditional independence of the encouragement (Rosenbaum and Rubin 1983). Estimating equation (1) with weights gives the so called *doubly robust* estimator that ameliorates this sensitivity to misspecification (Hirano, Imbens, and Ridder 2003). Note that this model is akin to a Blinder-Oaxaca decomposition in which I interpret the *unexplained* part as treatment effect heterogeneity (See e.g. Fortin, Lemieux, and Firpo (2011) and Ding, Feller, and Miratrix (2019) for more details and alternative methods).

The covariates include cohorts, unemployment status, number of children, years receiving RSA, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. Although it would be safer to let a machine-learning algorithm select the most relevant covariates, the dataset is small for a generalised random forest, and double-lasso may end-up creating omitted variable bias (Wüthrich and Zhu 2023). In this analysis, I chose simplicity, interpretability and consistency with other analyses although such analysis may be implemented in future revisions of the paper.

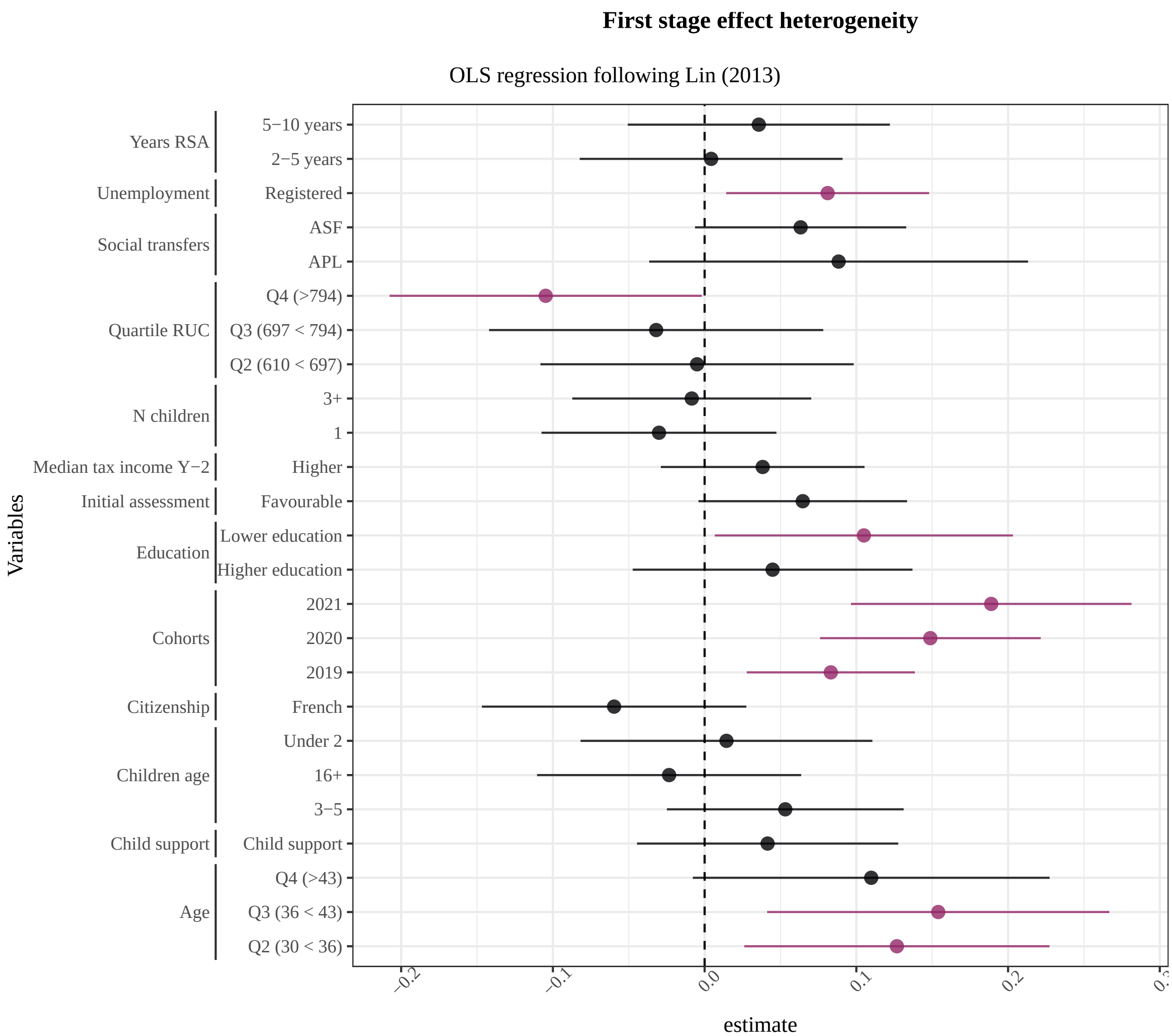
Higher participation among the least educated and poorest single mothers in their thirties Estimates of the $\hat{\boldsymbol{\beta}}$ coefficients are presented in figure 2 and confirm the previous observations on the heterogeneity w.r.t cohorts. The other results help defining a typical profile of compliers. First, those registered at the Employment agency and with favourable initial assessment are about 8 pp more likely to participate, with 95% pointwise confidence intervals between 1 and 15pp. Second, participation is higher for those with less than a high school degree (+ 10 pp) and the poorest at the time of random assignment. The 25 % with highest income are indeed -10pp less likely to participate than the 25% poorest. Age is also predicting higher participation for single mothers in their thirties, implying a lower participation among the youngest. Overall, these estimates shows that the programme attracted middle-aged mothers among the poorest and least educated of the sample. The larger take-up from those registered at the Employment agency may capture part of the latent *fear* of sanction or reflect a closer relationship with the labour market. In the control group, those registered at the Employment agency have a much higher job finding rate than unregistered mothers, as seen in Figure 1. In the absence of intervention, registered single mothers are more

[34] The instrument-propensity score is simply the proportion of the encouragement group in each block and weights: $w_{jc} = \frac{Z_{ijc}}{\hat{q}_{jc}} + \frac{(1-Z_{ijc})}{1-\hat{q}_{jc}}$

likely to find a job, and their employment level after the period of the programme is the highest of all the compared groups.

Moreover, those who receive the family support allowance (ASF) and with favourable initial assessment are slightly more likely to participate although I cannot exclude 0 from the 95% CI. ASF is the allowance paid when the other parent is unable to pay child support or while waiting for a judge to set the amount. Those who do note receive it either receive child support from the other parent or did not go through the administrative process of setting child support (Périvier and Pucci 2019). The other plausible covariates I introduce do not predict participation. In particular, neither the number of year receiving RSA, nor the number of children predict participation although they predict outcomes.

Figure 2: Estimating the heterogeneous effect of encouragement by baseline characteristics



Sources: ALLSTAT, observations one month before random assignment. Cohorts 2018 to 2021.
Notes: linear regression of participation on encouragement following Lin (2013): I regress participation on encouragement, demeaned covariates and the interaction of both. I use inverse probability weighting to account for the design.
I only present the coefficients of the interactions. The effect of encouragement and coefficients of the level variables are not shown (and very close to 0 by virtue of random assignment).
Point−wise 95 % confidence interval based on cluster−heteroskedasticity robust standard error with no degree of freedom adjustment (CR0).
Pink coefficients indicates coefficients that exclude 0 from the 95% CI.

III.4 Evolution of the main outcomes over the period

Figure 3 illustrates the average employment levels over time, while Figure B.16 in the Appendix depicts the evolution of disposable income per consumption unit. Both figures include means with point-wise 95% confidence intervals and a spline smoothing by encouragement status and cohort. In addition, I plot the timing of the reform increasing in-work benefits in the first quarter of 2019 and the three lock-downs of the Covid-19 pandemic. These four panels offer a clear overview of the employment trends. While there are small variations between cohorts, there are no significant differences between encouragement and control groups after the end of the programme and no sign of improvement for earlier cohorts.

Before random assignment, we observe some Ashenfelter's dips (J. Heckman et al. 1998) with downward pre-trends and the lowest employment rate on the quarter before sampling. Cohorts 2018 and 2019 are balanced at baseline, but not cohort 2020 and 2021 for which differences between groups are reversed. However, these imbalances are constant for at least two years and are differenced-out in the estimations. The time span between random assignment and the beginning of the programme varies between cohorts, letting the control group get ahead of the encouragement group. This is particularly true for cohort 2019 and 2020, but not at all for 2018, whose recruiting period was the shortest. For the 2019 cohort, the PA reform occurred right in the middle the recruiting period, probably contributing to the observed lock-in effect.

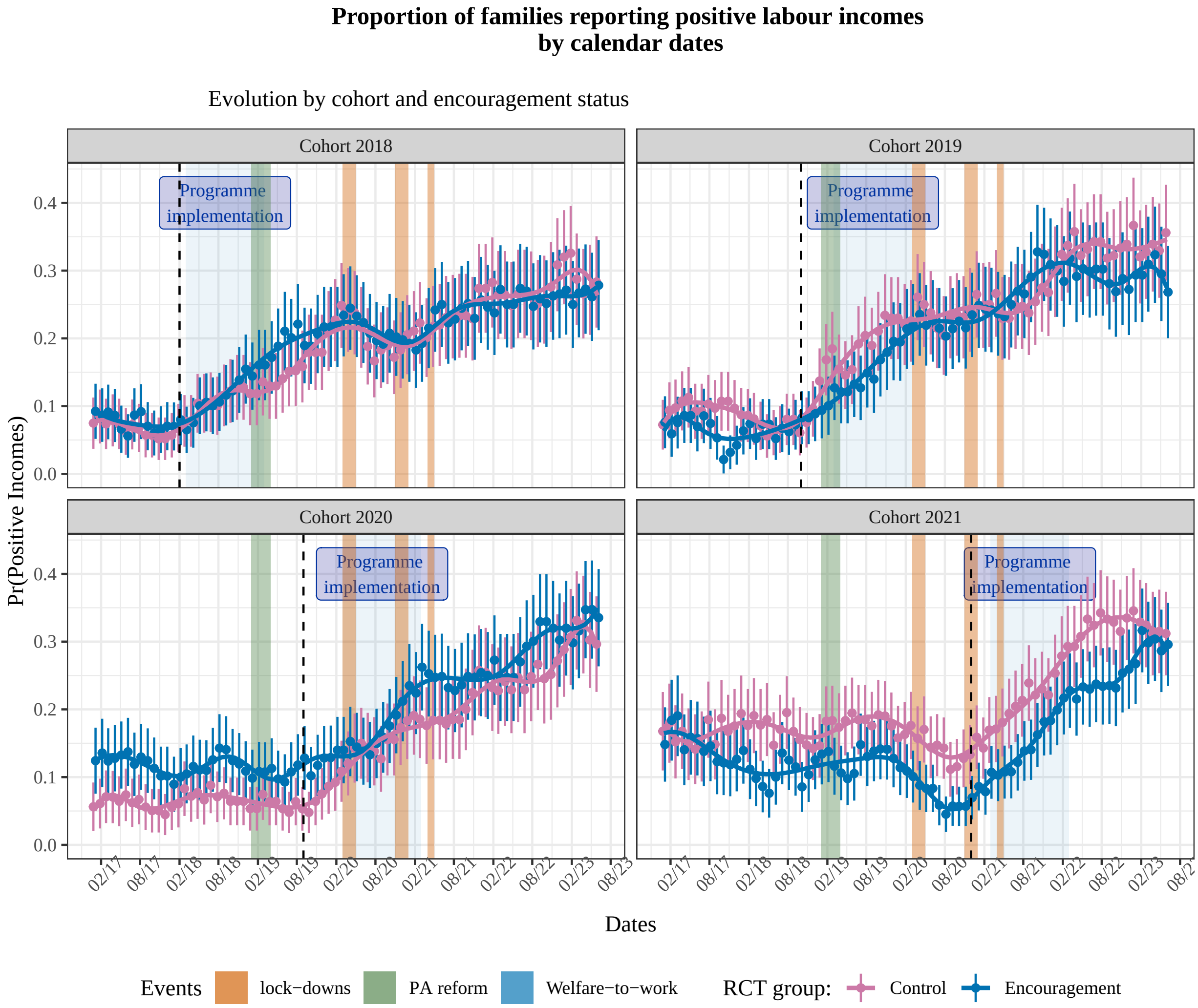
The first quarter of the training remains stable in all quarters while the qualitative analysis reports that this period is crucial in boosting motivation and self confidence, and start a positive dynamic. It, start showing up on the second quarter but the employment growth in the control group is still higher.

In the second half of the programme, the job finding rates of encouragement groups notably increase, even for the 2020 cohort in the midst of the pandemic. For the 2018 cohort, there is a short-lived positive difference favouring the encouragement group, which coincide with the reform of in-work benefits. But the control group caught-up 6 months later and employment rates are the same for the rest of the observation window. For the 2019 cohort, the programme ended right at the beginning of the pandemic with no difference between groups. Employment dynamics froze until the last lock-down where job finding rates started to increase again for both groups. If anything, the encouragement group has a slightly lower employment level than the control group. Finally, Cohort 2020 and 2021 ends-up with the same employment levels in both groups at the end of the programme. Note that ultimately, roughly 1/3 of the sample has positive labour income. This is true for all cohorts and randomised groups. The 2018 cohorts is the slowest to converge, and the pace seems to increase in recent cohorts despite the pandemic. Whatever the duration since we sampled the cohort, about 2/3 of each cohort remains without a job.

The dynamics of disposable incomes per consumption units presented in Figure B.16 in the Appendix is tightly linked to that of employment, although it shows meaningful differences. The encouragement group of the 2018 cohort is significantly poorer than the control group in the long run. The encouragement group of the 2019 cohort was significantly richer than the control group before random assignment and this difference correspond to the gap we observe at the end of the period. There are strong lock-in effects for all groups. The 2021 encouragement group was significantly poorer than the control group and the initial gap remains after the end of the programme.

Overall, these descriptive statistics show that on average, these households are and remain poor, with little evidence of an effect in intention-to-treat. I now describe my empirical strategy.

Figure 3: Average proportion of families reporting positive work income over the period



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Points indicate the sample mean by date and randomisation group. Error bars indicates pointwise 95 % CI using .975 quantile of a normal distribution, sample size and variance.
Smooth lines are estimated using splines with 1/6 of the number of dates as degree of freedom.

IV Identification and estimation strategies

I mostly focus my analysis on four target parameters:

1) Average effects of encouragement on participation (First stage)
2) Average characteristics of compliers and never-takers
3) Dynamic and aggregated average intention-to-treat effects
4) Dynamic and aggregated average treatment effects on the treated.

Being the only experimental evaluation of such programme in France, I also estimate the dynamic average treatment effect on the treated in an event study design as if I had no randomised experiment. Comparing these estimates with those obtained with the instrumental variable strategy sheds light on the importance of selection bias.

Identification stems from random assignment of encouragement within blocks of single parents. At each point in time, there are control units of the same cohort and from the same block whose observed outcomes serve as counterfactual. Therefore, all well conducted comparisons between the encouragement and control groups have a causal interpretation. In particular, block-random assignment gives a causal interpretation to the first stage and comparisons of outcomes are interpreted as *intention-to-treat* parameters (S. Athey and Imbens 2017b). Then, assuming encouragement has no other effects on outcomes but through its effect on participation, an instrumental variable strategy yields the effects on compliers, which in this setting with one-sided compliance, can be interpreted as *treatment effect on the treated* (Frölich and Melly 2013). In practice, the main challenges are i) to aggregate treatment effects across blocks, cohorts and time without imposing restrictions on heterogeneity and ii) compute appropriate standard errors and correct for simultaneous inference.

IV.1 First stage

It is well known that the average causal effect of the encouragement on participation across cohorts can simply be obtained by regressing participation on encouragement and block x cohort fixed effects using OLS and cluster-robust standard errors[35] (Negi and Wooldridge 2021; Alberto Abadie et al. 2022; C. de Chaisemartin and Ramirez-Cuellar 2022).

By the Frisch-Waugh-Lovell (FWL) theorem, this regression is equivalent to regressing 1) participation on block fixed effects 2) encouragement on block fixed effects and 3) residual of the first one on the second. The first auxiliary regression removes the average within blocks, therefore averaging control and encouragement together. Since there are no treated unit in the control group, their predicted values are out of the [0;1]. This model is therefore far from the conditional expectation function, as illustrated in Figure E.28 in the Appendix. While this misspecification is not a problem to estimate the average effect on participation, it has important consequences when the model serves as a first stage of a TSLS.

First stage for the second stage Since the projection matrix of the first stage is part of the solution of the TSLS system, the model is biased and does not recover the LATE. The recent work of Blandhol et al. (2022) revisits the use of TSLS to estimate the LATE and shows that the only specifications that have a LATE interpretation are i) "saturated" specifications that control for covariates non-parametrically or ii) models with restrictive parametric assumptions. Moreover, the fully saturated specification is more sensitive to finite-sample bias than estimates of just-identified models, especially in small blocks with low compliance[36]. Tübbicke (2023) compares different IV estimators and confirms that TSLS are likely biased when covariates predict the instrument and when groups are of different sizes.

In contrast, the recent work by Borusyak, Hull, and Jaravel (2022) provides an alternative but simple fix to this problem: de-mean the instrument by the block-specific instrument propensity score. The latter only depends on the proportion of encouraged households in each block that represents a *type* of household. The instrument propensity score is the assignment probability for each type and is given by the design, thus satisfying Assumption II of Borusyak, Hull, and Jaravel (2022). Let $\tilde{Z}_{ijc} = Z_{ijc} - \hat{q}_{ijc}$ be the "re-centred" offer instrument that subtracts the instrument propensity score q_{ij} from the encouragement indicator in block Z_{ij}. To estimate q_{ijc} I simply run a probit regression of Z on block-cohort fixed effects and use its predicted probability $\hat{q}_{ijc}$. I then estimate the following regression using OLS:

$$D_{ijc} = \sum_c \sum_j \alpha_{jc} A_{ijc} + \pi \tilde{Z}_{ijc} + \varepsilon_{ijc}, \quad i \subset \mathcal{P}_{\{\cdot, m=c\}} \tag{2}$$

Where D_{ijc} is a dummy for participation for parent i of block j from cohort c, A_{ijc} are block fixed effects (types) and $\tilde{Z}_{ijc}$ the demeaned instrument for encouragement. I estimate this model on the month of random assignment. To further improve the validity of this model, I weight observations by the inverse of the instrument propensity score (IPW) to obtain a so-called doubly-robust estimand. The weights used are $w_{IPW} = \frac{Z_{ijc}}{\hat{q}_{ijc}} + \frac{1-Z_{ijc}}{1-\hat{q}_{ijc}}$

[35] I discuss and justify the choice of clustering in a dedicated section below.
[36] The coefficient of the fully saturated regression are represented in Figure E.29 in the Appendix and the prediction in figure E.28.

To gain intuition, note that by the FWL Theorem, the estimand is unchanged if one includes q_{ijc} as control and the re-centred instrument $\tilde{Z}_{ijc}$i is replaced with the unadjusted offer Z_{ijc}. Since the propensity score is given by design, the main theorem of Rosenbaum and Rubin (1983) applies and this ensures unbiasedness of the IPW regression (Tan 2010; Cattaneo 2010).

This model can accommodate additional covariates to improve precision, provided they are centred. Negi and Wooldridge (2021) show that there are no gain in a full regression adjustment i.e. interacting the treatment with demeaned covariates, when Z is balanced, which is the case here as shown in table B.6.

While these adjustment have little impacts on the estimation of the average effects as the propensity scores are close to .5 and only differ because of roundings in the assignment, they matter for the causal interpretation of TSLS, as I know discuss.

IV.2 Characterising the full distribution of compliers' potential outcomes and attributes

Random assignment reveals potential participation and help identification of the effect on compliers (J. D. Angrist and Imbens 1995). To introduce notations for potential participation and other variables, let $V^k(d) = V^k(D(z)) \equiv V^k(d,z)$ denote the potential values of a variable V as function of participation and encouragement.

By design, $V^k(1,0) = \emptyset$ for there are no always takers. $V^k(1,1)$ is the revealed potential value of V for treated compliers and $V^k(0,0)$ the revealed potential value for the control group.

The key assumption for instrumental variable is the **exclusion restriction** where V is replaced by the set of outcomes Y:

Hypothesis 0.1 (Exclusion)**.**

$$Y^k(d,z) = Y^k(d,z') = Y(d) \;\forall\; z,\; z',\; d$$

which rule out an effect of encouragement, the instrument Z_{ijc}, on the variable of interest Y^k_{ijct} but through its effect on participation. In particular it means that for the never takers, $Y^k_{ijct}(D_{ijc}(0),0) = Y^k_{ijct}(D_{ijc}(0),1) = Y^k_{ijct}(D_{ijc}(0))$, which is actually the only restriction on potential outcomes we need (S. Athey and Imbens 2017b). In words, the encouragement has no effect on outcomes on its own, it only affects them through increased take-up in Reliance and the intention-to-treat effect is only driven by the treatment effect on compliers.

The exclusion restriction is a strong hypothesis and cannot formally be tested. In particular, it means that never-takers have no other reaction to the encouragement but to discard the invitation. Arguably, the encouragement might make people feel a little more scrutinised by the administration and threatened by welfare reduction if they do not comply. However the most effective way to avoid sanctions is to participate in the meeting which supports our hypothesis. Furthermore, this population receives letters from social services all the time and our invitation was no more threatening than any other. They are used to update their situation every quarters. I therefore assume the exclusion restriction holds.

The first stage allows to estimate the average proportion of compliers across cohorts $\pi = \Pr\left(D_{1ijc} > D_{0ijc}\right) = E\left(D_{ijc} \mid Z_{ijc}=1\right) - E\left(D_{ijc} \mid Z_{ijc}=0\right)$. The problem is that we don't know who the compliers are. When we see a particular observation among non-participants, it can be either never-taker or complier. But it can only be complier among participant because of one-sided non compliance. A well-known result discussed in A. Abadie (2003) and Frölich and Melly (2013) is that one can identify the distribution of the characteristics of the compliers. In practice, A. Abadie (2003) proposes a simple 2SLS procedure for characterising compliers, described by:

$$g\left(X_{ijc}, Y_{ijc}\right) \times \mathbb{1}(D_{ijc}=d) = \sum_c \sum_j \mathbf{A}'_{ijc}\alpha_{jc} + \gamma_d \mathbb{1}(D_{ijc}=d) + \mu_{ijc} \quad d \in \{0,1\} \tag{3a}$$

$$\mathbb{1}(D_{ijc}=d) = \sum_c \sum_j \mathbf{A}'_{ijc}\alpha_{jc} + \tilde{Z}_{ijc}\pi + \epsilon_{ic} \tag{3b}$$

where $g\left(X_{ijc}, Y_{ijc}\right)$ is any function of family baseline characteristics X_{ijc} and outcomes Y_{ijc}. Setting $d = 0$ means I use Z_{ijc} to instrument $(1 - D_{ijc})$ in an equation with $g\left(X_{ijc}, Y_{ijc}\right)(1 - D_{ijc})$ as the outcome.

Borusyak, Hull, and Jaravel (2022) show how such specifications identify weighted averages of conditional-on-block IV coefficients[37]. In particular, the IV estimand is given by:

$$\gamma_d = \int w_d(t)\gamma_d(t)dF_j(t) \tag{4}$$

where $F_j(\cdot)$ gives the distribution[38] of blocks j and $\gamma_d(t)$ the block specific average of compliers.

This estimation thus recovers average characteristics of treated and untreated compliers:

$$\gamma_d = E\left[g\left(X_{ijc}, Y_{ijc}(d)\right) \mid D_{ijc}(1) > D_{ijc}(0)\right], \quad d \in \{0,1\} \tag{5}$$

By setting $g\left(X_i, Y_i(d)\right) = X_i$, the IV procedure produces the average of any predetermined covariate X_{ijc} for compliers within a block, and I estimate never-taker means by regressing $X_{ijc}(1 - D_{ijc})Z_{ijc}$ on $(1 - D_{ijc})\tilde{Z}_{ijc}$ with block fixed effect.

With this transformation, I can also estimate the distribution of the missing potential outcomes for compliers. By setting $g\left(X_i, Y_i(d)\right) = \mathbb{1}(Y_i \leq y)$ for a constant y and each value of d, I obtain the complier cumulative distribution functions of $Y_i(1)$ and $Y_i(0)$ evaluated at y.

To estimate the potential densities for treated and untreated compliers, Abdulkadiroğlu, Pathak, and Walters (2018) propose to use a symmetric kernel function $g\left(X_i, Y_i(d)\right) = \frac{1}{h}K(\frac{Y_i - y}{h})$ in equation (3a) to estimate the density of the potential outcomes for compliers at $Y = y$. $K(\cdot)$ is a symmetric kernel function maximised at zero and h is a bandwidth that shrinks to zero asymptotically. I follow Abdulkadiroğlu, Pathak, and Walters (2018) and evaluate complier densities at a grid of 100 points with kernel bandwidth defined as $h = 1.06 \times N^{\frac{-1}{5}}\sigma_d$, where N is the sample size and σ_d is the standard deviation of the potential outcome.

IV.3 Dynamic treatment effects

I follow a design-based approach and consider two key parameters of interests for every relative month since random assignment:

- The average intention-to-treat effect of the programme
- The average treatment effect on the treated

Both are weighted average of block x cohort x time average effects. There are thus three key challenges. First, we need estimates of the block x cohort x time average effects, ITT and LATE. Second, we need to properly aggregate theme and third, we need to compute standard errors and account for simultaneous error.

With panel data, a natural way of estimating the dynamic treatment effects is to estimate cohort-time difference-in-differences and aggregate them in an event study. However, it is now well known that the staggered adoption of treatments with time-varying dynamic effects can cause sever issues in regression models without appropriate corrections or alternative estimands[39].

There are two design-based estimands that directly allow to estimate the event study parameters. My preferred models are based on the so called "stacked regression" which I present in details in the next sub-section. The other estimand is a modified version of Callaway and Sant'Anna (2020a) that requires separate estimates for each cohort and aggregation using the share of each cohort as weights. For intention-to-treat estimations, randomisation conditional on blocks ensure parallel trends and both methods yields causal estimates.

[37] See the chapter by J. Angrist, Hull, and Walters (2023) with a very clear presentation and application to school choice and the recent review by Borusyak, Berkeley, and Hull (2023) for more intuition on the theoretical results and review of other applications.
[38] Here, the distribution of weights is simply given by the share of each block in the sample.
[39] See e.g. Roth et al. (2021) for a recent review

A) Main model for dynamic intention-to-treat analysis:

The stacked-regressions consider each cohort as a sub-experiment and imposes a fixed event time windows relative to the month before random assignment. Let $m = t - c$ denote the relative time since randomisation such that $m \in \mathcal{M} =$ [-13,30], the upper limit being the last observation of the 2021 cohort. I estimate the following equations using OLS:

$$Y_{ijcm} = \sum_m \sum_c \sum_j \alpha_{jcm} A_{jc} + \sum_m \beta_m \tilde{Z}_{ijc} \times \mathbb{1}(t = c + m) + \sum_m \mathbf{X}' \boldsymbol{\rho}_\mathbf{m} + \varepsilon_{ijcm} \tag{6}$$

Where Y_{ijcm} is the outcome of household i in block j from cohort c, observed $m = t - c$ months since random assignment. A_{jc} are block $\times$ cohort fixed effects and $\tilde{Z}_{ijc}$ is the demeaned encouragement variable. I include a set of baseline covariates in a matrix $\mathbf{X}$ with a set of coefficient $\boldsymbol{\rho}_\mathbf{m}$ for each period (see below). Like before, I reweight observation by the inverse instrument propensity score to obtain doubly robust estimates. The interaction of all block $\times$ cohort with relative month dummies and the joint estimation of all ITTs yields a fully saturated regression. It estimates a treatment variance weighted average treatment effect of block specific intention-to-treat. One way to think about this regression is as a way of estimating all of the $ITT(c,t)$ parameters and then immediately aggregating them into a single set of event time parameters. Intuitively, it is equivalent to a regression of the long difference between month m and the reference month on the demeaned treatment variable and block cohort fixed effects restricted on the window of observations where cohorts are m months since random assignment. Each of these equations are stacked by interacting all right hand side parameters with relative time dummies. This method uses what looks like a typical TWFE regression estimate, but because of the structure of the data, it only incorporates clean controls. Notice that this model is similar to the event-study estimand of Sun and Abraham (2020). The only difference is that I use interactions of blocks instead of cohorts to better reflect the experimental design.

Adding covariates to the model The main motivation for adjusting for covariates is that the precision of the estimated average treatment effect can be improved if the covariates are sufficiently predictive of the outcome (Lin 2013). Moreover, I include one dummy for individuals resampled and one for those encouraged in the 2022 cohort to allow a specific trajectory for this subsample. Randomisation ensures there is no correlation in the population between Z_{ijc} and the covariates $\mathbf{X}_i$, which is sufficient for the lack of bias from including or excluding the covariates (S. Athey and Imbens 2017b). I follow Negi and Wooldridge (2021)'s recommendation and use pooled regression since $\mathbb{E}[p_{jc} = .5]$ in the random assignment case is precisely the condition that implies no efficiency gain from full RA even when there is arbitrary heterogeneity in the treatment effects. In particular, including baseline outcomes (centred) may increase precision while leaving the p-limit of the ITT coefficients unchanged. In practice, I add covariates interacted with relative time dummies to have a coefficient for each relative time period, allowing specific relative date effect of each covariate.

Inference Because I expect heterogeneous treatment effects across blocks, I use cluster-robust standard errors[40] (CR0 typically used in Stata) (Liang and Zeger 1986) at the block-cohort level following the recommendations of Alberto Abadie et al. (2022). It is worth noting that these corrections for clustering are asymptotically unbiased and converge to the *super population* parameter - the latter being rather undefined. A more design-based inference could be derived using results from Alberto Abadie et al. (2020). However, since we have both random sampling and random assignment, there are not much gains to expect and we can rely on the usual asymptotics. A last choice is about the degree of freedom adjustments and C. de Chaisemartin and Ramirez-Cuellar (2022) conveniently summarize their recommendations for practitioners. In all models, I do not adjust the degree of freedoms following their recommendations for models with fixed effects, clustering and blocks with less than 10 units.

Because the analysis focuses on the dynamics of the treatment effects at different points in time, the chances of type-II error increase and the test statistics based on point-estimate standard error will over-reject the null hypothesis of no treatment effect. Instead, I consider testing K null hypotheses[41] $H_0^1, \ldots, H_0^K$ individually and require that the family-wise error rate, *i.e.*, the probability of falsely rejecting at least one true null hypothesis, is bounded by the

[40] Cluster levels are equivalent to a regression on data collapsed by block $\times$ cohort averaging the full set of block specific treatment effects.

[41] Where K can refer to m or other sets of joint hypotheses.

nominal significance level $\alpha = .05$. In what follows I use adjusted p-values to describe the decision rules. Adjusted p-values are defined as the smallest significance level for which one still rejects an individual hypothesis H_0^j, given a particular multiple test procedure (Hothorn, Bretz, and Westfall 2008). By construction, I can reject an individual null hypothesis $H_0^k, k = 1, \ldots, K$, whenever the associated adjusted p-value is less than or equal to the pre-specified significance level α, i.e., $p_k \leq \alpha$. Most specifications consider the estimation of the ITTs over 12 months before randomisation and up to 30 months after with 4 cohorts or 45 months with 3 cohorts. Simultaneous inference for that many parameters costs a lot of power. To gain precision, I define $S \equiv S(m) = \lfloor m/6 \rfloor$ the relative semester since randomisation to average over 6 months and estimate average effects over each period.

The parsimonious aggregated models estimate the following equation:

$$Y_{ijcm} = \sum_m \sum_c \sum_j \alpha_{jcm} A_{jc} + \sum_s \beta_s S(m) \times \tilde{Z}_{ijc} + \varepsilon_{ijcm} \tag{7}$$

Like before, the Frisch-Waugh Lowell theorem makes clear what this equation identifies. The first part of the right hand side of the equation removes any average differences *between* blocks at every relative month and leaves the *within-block* variation. It also removes differences in the share of encouraged households across blocks - which is akin to subtracting the block-specific propensity score of encouragement from the treatment dummy. I come back to this point in the description of the instrumental variable estimation strategy. The estimations $\hat{\beta}_s$ of this equations therefore delivers a variance weighted average of block-cohort-relative months specific treatment effects. As S. Athey and Imbens (2017b) note, in general, this parameter is not an unbiased estimand of the ITT over the permutation distribution of assignment.

B) Adaptation of Callaway and Sant'Anna (2020b)

Among the new methods for difference-in-differences with staggered adoption, the method proposed by Callaway and Sant'Anna (2020b) is very well fitted for this setting. Their event-study estimators:

1. Estimate a generalised propensity score based on the covariates
2. Estimate every 2x2 difference-in-differences between the reference point and another date between treated and untreated individuals in the cells with inverse-propensity score weightings similar to Alberto Abadie (2005). These are group-time treatment effects (GTTEs)
3. Aggregate the group-time GTTEs weighting them by the relative size of each group in the sample to compute the event-study coefficients.
4. Estimate standard errors accounting for multiple testing using wild cluster bootstrap.

The main difference with the previous method is a new estimation of the propensity score including covariates, the use of the doubly robust estimator by Sant'Anna and Zhao (2020) and simultaneous inference using a wild cluster bootstrap. The doubly robust estimator estimates the sample analog of the following equation:

$$\text{ITT}(c,t) = \mathbb{E}\left[\left(\frac{Z_{ic}}{\mathbb{E}[Z_{ic}]} - \frac{\frac{p_c((X))(1-Z_{ic})}{1-p_c((X))}}{\mathbb{E}[\frac{p_c((X))(1-Z_{ic})}{1-p_c((X))}]} \right) (Y_t - Y_{c-1}) \right] \quad \text{under conditional parallel trend} \tag{8}$$

However, assumption 2 in Callaway and Sant'Anna (2020b) imposes that each unit i is randomly drawn from a large population of interest. Here, the sampling probability is the same within cohort, but varies across. That means that using the CS-2020 estimator on the full sample would use observations of the untreated units of other cohorts in the comparison group. A simple way to correct this is to run steps 1 and 2 separately for each cohort panels, and then aggregate the parameters in an event study using a similar framework. This is what I call the "modified CS-2020"[42]. This estimand is doubly robust and does not restrict heterogeneity.

[42] I thank Pedro Sant'Anna for his help in adapting the DID package. The R code for this function is available upon request

The aggregate estimates are obtained by weighting each group time treatment effects by the following weights:

$$\theta_{es}(m) = \sum_{c} \mathbf{1}\{t = c+m\} P(C = c \mid t = c+m) TT(c, c+m) \tag{9}$$

In words, each cohort-time average treatment effects is weighted by its relative size in the sample of observations at time $t = c+m$. Substituting Z by D in equation (8), I could obtain an estimate of the average treatment effect on the treated assuming conditional parallel trend with an appropriate control group. I estimate these models to analyse endogenous selection in section VI.

IV.4 Treatment effects on the treated

To estimate the average treatment effect on the treated using the encouragement as an instrument for participation, I simply estimate TSLS system of stacked regressions.

$$Y_{ijm} = \sum_{m \neq -1} \sum_{c} \sum_{j} \mathbf{A}'_{ijm} \beta_{jcm} + \sum_{m \neq -1} \delta_m D_{ijc} \times \mathbb{1}(t = c+m) + \mu_{ic} \tag{10}$$

$$D_{ijc} \times \mathbb{1}(t = c+m) = \sum_{m \neq -1} \sum_{c} \sum_{j} \mathbf{A}'_{ijc} \alpha_{jcm} + \sum_{m \neq -1} \pi_m \tilde{Z}_{ijc} \times \mathbb{1}(t = c+m) + \epsilon_{ic} \tag{11}$$

Each block x cohort x relative time fixed effects instrument themselves in the second equation while participation at every month is instrumented by the interaction of the demeaned instrument with the corresponding relative month dummy.

Because of one-sided non compliance, these parameters correspond to the average treatment effect on the treated (Frölich and Melly 2013). This result on shift share IV immediately applies to block-random encouragement design[43].

Like the previous models, I adjust standard errors for clustering at the block $\times$ cohort level. In some settings, $m = f(t-c)$ can aggregate several months in a single parameter. The estimator is then a positively weighted average of the LATEs over the period (Mogstad, Torgovitsky, and Walters 2021). Just like for the ITT, the repetition of statistical tests increases the chance of false positive and I use the same method as with the ITT to correct for the family-wise error rate and use a 5% adjusted p-value to define 95% adjusted confidence intervals.

Remarks Equations (10) and (11) are actually an event study version of a *fuzzy difference-in-differences* (FDID) estimator that recovers the LATE if one further assumes a stable percentage of treated unit in the control group over time (C. D. Chaisemartin and D'Haultfoeuille 2017, Theorem 1) which is ensured by design.

Adding covariates to equations (10) and (11) instrumenting themselves can improve precision but has two important consequences: first it imposes constant treatment effects and second, it loses the LATE interpretation and may introduce bias through the linear projection of the first stage that may be further away from the true conditional effect on participation. Adding covariates to a linear equation and interacting them with the treatment indicator seems like a natural way to account for non-random assignment of the instrument while allowing for heterogeneous treatment effects. Unfortunately, no result implies that this procedure generally uncovers the LATE (Słoczyński, Uysal, and Wooldridge 2022). Nevertheless, this model with covariates can be informative. If we observe differences between the LATE (as estimated without covariates) and estimates assuming constant treatment effect and covariate-specific time effects, then the constant treatment effect assumption is likely false and motivates the analysis of heterogeneous treatment effect.

[43] This result has been applied in the school choice literature (Joshua D. Angrist et al. 2017; Abdulkadiroğlu, Pathak, and Walters 2018) and is central to my work on daycare assignment (Combe and Heim 2024). All the future work on evaluation lay on this set of recent results.

V Does it work ? Main results on labour and poverty

This section presents the main results of this evaluation: the effects of the programme on labour market participation and poverty. I first start with the dynamics of the treatment effects on labour market participation, disposable income per consumption units and risk of poverty. I also estimate the distribution of potential incomes for treated and untreated compliers to gauge distributional effects. I interpret these results and complete the picture with supplementary estimates in Appendix C. In brief, I find that the programme generates a strong lock-in effects that slows employment, reduces disposable incomes and slows exit from poverty. These negative effects fade-out by the end of the programme. In the post-treatment period, there are no average effects of the programme on employment and disposable incomes and 90 % of the sample remains in poverty. In the year following the programme, I cannot reject that the distribution of disposable incomes are the same.

V.1 Lock-in and no average post-treatment effects on employment or disposable incomes

Figure 4 presents the main results on intention-to-treat effects on labour market participation aggregating the four cohorts. The top panel shows the average levels of employment in encouraged and control groups while the second panel shows the estimates of equation (6): the dynamic intention-to-treat effects, with or without covariates. I also present estimates using the modified version of the estimator of Callaway and Sant'Anna (2020a), estimating all treatment effects of equation (8) and aggregating them with weights of equation (9). Results are almost identical although the doubly robust estimator is both more precise and slightly more negative during the lock-in period.

Longer recruiting processes impose large opportunity costs to participants Before random assignment, both groups had a similar level of employment, which slowly decreased. This pattern is akin to an Ashenfelter dip (J. J. Heckman and Smith 1999) although there are no differences between groups. In the first half of the programme, job-finding rates are higher in the control groups while employment dynamics of the encouragement group only fasten after the first half of the programme. Lock-in effects are very common in the literature. During this period, job-search intensity may be lowered because there is less time to search for a job, and participants may want to complete an on-going skill-enhancing activity. A common and intuitive result often reported is the the longer the recruitment or the programme lasts, the larger the lock-in effect[44]. Over the years, we have extended the period between the draw and entry into the programme to maximise participation[45]. And while larger shares of the targeted groups signed-in, they also lost job opportunities waiting for the programme to start, or to benefit from it.

The qualitative report explains that the initial three months of the programme facilitated a rapid shift in momentum, boosting confidence and self-esteem relatively quickly. The group dynamic and "lever effect" of collective activities helped individuals overcome isolation. For the *slowed down* (mothers with access to resources but needing support to regain momentum), the pathway to stable employment or swift enrollment in training became viable. The *motivated* participants (mothers with limited social resources, often with migrant backgrounds, but determined and with some work experience) were poised for short-term employment opportunities, typically in sectors requiring minimal training and with higher recruitment rates.

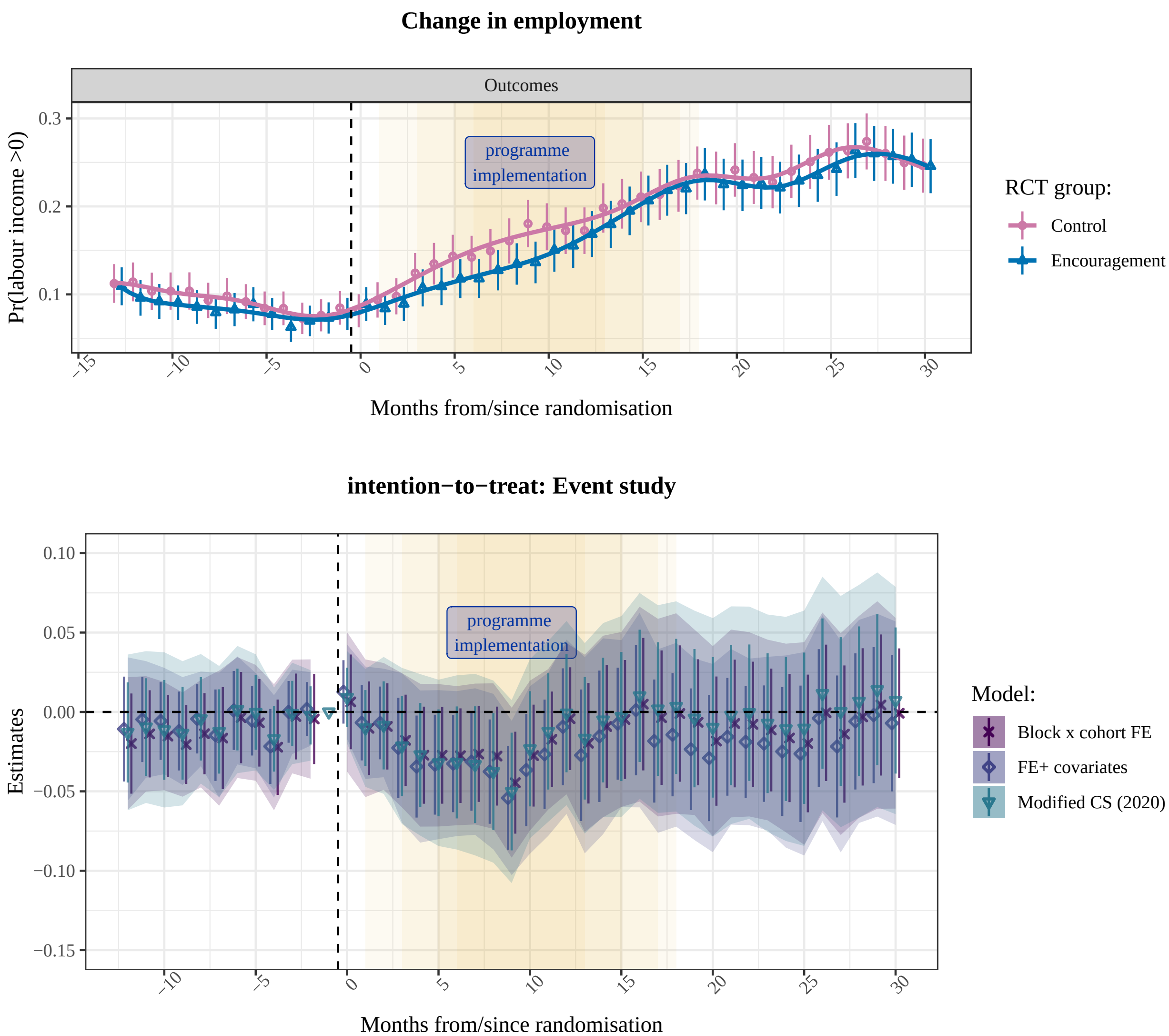
These benefits show up on the job-finding rate after the 10^{th} month since randomisation. In one quarter, the encouragement group has gained on the the control group and after the programme, employment levels are virtually the same in both groups. Conversely, the qualitative evaluation also explains that isolated individuals with limited work experience and reduced autonomy (called the *excluded*), and those with active social networks but not actively seeking employment due to caregiving responsibilities, health issues, etc. (called the *hindered*) faced a longer timeline for employment reintegration, as their focus was not on immediate employment but rather on formulating a professional plan.

In the end, there are no average intention-to-treat effect after the end of the programme and the employment level is about 24.4% from 18 to 30 months after random assignment. In Figure E.32, I show the same estimations for

[44] See for instance Wunsch (2016) or Filges et al. (2015)

[45] An other reason is that it takes a lot of time to check every file before random assignment and managers did not want social workers overstaffed.

Figure 4: Dynamic Effects of the programme on employment



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for positive individual labour income.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS or modified Callaway Sant' Anna (2020).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level .
− Shades indicates 95%CI adjusting for the FWER uing the Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

the three first cohort up to 45 months. Results are less precise but qualitatively similar. They show no sign of improvement on the 15 additional months of observation.

Imprecise estimates still excluding policymakers's minimal expected effects Simultaneous confidence intervals barely exclude 0 on the 9^{th} month where the employment dynamics shift. Part of this imprecision and lack of statistical power - despite having one additional and larger cohort than initially - comes from unanticipated and large treatment effect variations across cohorts, and larger overall variance especially in the control group. The higher variability comes from difference macro-economic conditions while participants in the programme had rather similar trajectories[46].

To get a broader picture and estimate more precise parameters, I average the treatment effects over periods of 6 months and report the estimates in table C.7 in the Appendix. These estimates confirm the previous results and the-3.65pp of the ITT point estimate with covariates during the lock-in period correspond to -22.3 % of the employment level in the control group. Accounting for simultaneous inference, there are 95 % chance that the average ITT is between -7.3pp and 0 pp. The adjusted p-value for an effect different from 0 is 0.051. Using instrumental variables, The lock-in effect on employment for participant is -9.3 pp, with simultaneous 95% between -18.4pp and -0.3pp.

From 24 to 30 months after random assignment, the average treatment effect on the treated is -3.6pp, and I can rule out an average positive effect on the treated as high as 8.8 percentage point with 95 % confidence, thus excluding the minimal policymakers' target of 10 percentage points.

Lost opportunities, lost incomes Figure C.17 in the appendix completes the picture by presenting the average intention-to-treat effect on disposable incomes per consumption unit. This measure is directly computed by the National family allowance fund and accounts for family size. I use the Consumer Price Index for the bottom income quintile for actualisation in 2015 €. Obviously, the observed patterns are similar to the previous results but confidence intervals are more informative, and others are worth noticing. First, the seasonality in the top panel essentially stems from the imputation of one-time-per-year social transfers[47] by the Family allowance fund. Nevertheless, this panel shows that in real term, the average disposable income was and remained well below the poverty line and single mothers in this sample were getting poorer and poorer. The employment divergence noted above can be observed in the change in average outcomes. Between 6 and 12 months since random assignment, the control group had 727 $€_{2015}$ of income per capita and the encouragement group's was -30 $€_{2015}$ lower, which represent -4.1% of the control group average. The negative effect on disposable income takes longer to fade but in the end, both groups have the same average disposable income.

I present the parsimonious estimations over a 6-month period in Table 2 and also include estimates of the average treatment effects on the treated (ATT) estimated and the average employment rate in the control group. The ATT is estimated instrumenting participation for each period by the demeaned encouragement interacted with period dummies and fixed effects. According to these estimates, participants lost an average of -76 $€_{2015}$ per month over the 6 first months of the programme. A back-of-the-envelop computation can give us the average income loss for compliers who did not get a job during the lock-in: If we divide this amount by the average treatment effect on employment discussed before (-9.3pp), we get 818. I discuss the effect of the programme on the intensive margin in Galitzine and Heim (2024).

This first set of results shows that the programme did not improve employment on average and caused participants to miss job opportunities during the lock-in period.

[46] See Figure 3 in section III to see the outcomes by encouragement, cohort and calendar dates.
[47] Christmas and back-to-school allowances.

Table 2: Aggregated effects of the programme on incomes per consumption units

	Mean control	OLS		TSLS	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	709.7*** (7.4)	-5.2 (7.1)	-5.4 (6.0)	-13.4 (18.2)	-14.0 (15.3)
	[691.1, 728.3]	[-23.1, 12.7]	[-20.8, 10.0]	[-59.5, 32.8]	[-53.4, 25.5]
	adj.p.val. = 0.000	*adj.p.val. = 0.923*	*adj.p.val. = 0.878*	*adj.p.val. = 0.924*	*adj.p.val. = 0.878*
[0 ; 6 [	704.6*** (5.7)	-15.4** (6.9)	-16.1*** (4.2)	-39.5** (17.9)	-41.3*** (10.8)
	[690.1, 719.0]	[-33.0, 2.1]	[-26.9, -5.3]	[-84.8, 5.8]	[-69.1, -13.5]
	adj.p.val. = 0.000	*adj.p.val. = 0.107*	*adj.p.val. = 0.001*	*adj.p.val. = 0.111*	*adj.p.val. = 0.001*
[6 ; 12 [	727.3*** (7.5)	-24.5** (9.3)	-29.8*** (8.6)	-62.8*** (23.6)	-76.4*** (21.9)
	[708.3, 746.4]	[-48.1, -0.9]	[-52.0, -7.6]	[-122.6, -3.0]	[-133.0, -19.9]
	adj.p.val. = 0.000	*adj.p.val. = 0.038*	*adj.p.val. = 0.003*	*adj.p.val. = 0.036*	*adj.p.val. = 0.003*
[12 ; 18 [	727.3*** (6.9)	-14.7* (8.5)	-18.7** (7.8)	-37.6* (21.3)	-47.9** (19.5)
	[709.8, 744.9]	[-36.2, 6.9]	[-38.9, 1.5]	[-91.6, 16.5]	[-98.2, 2.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.294*	*adj.p.val. = 0.083*	*adj.p.val. = 0.278*	*adj.p.val. = 0.069*
[18 ; 24 [	740.0*** (8.5)	-8.9 (8.4)	-13.1* (7.8)	-22.8 (21.0)	-33.3* (19.3)
	[718.6, 761.4]	[-30.1, 12.2]	[-33.1, 6.9]	[-75.9, 30.4]	[-83.1, 16.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.727*	*adj.p.val. = 0.357*	*adj.p.val. = 0.721*	*adj.p.val. = 0.338*
[24 ; 30 [	755.2*** (10.2)	-10.0 (12.3)	-12.1 (11.6)	-25.5 (31.1)	-30.9 (29.0)
	[729.6, 780.9]	[-41.2, 21.1]	[-41.9, 17.6]	[-104.2, 53.1]	[-105.7, 44.0]
	adj.p.val. = 0.000	adj.p.val. = 0.884	adj.p.val. = 0.793	adj.p.val. = 0.883	adj.p.val. = 0.793
Num.Obs.	28700	57927	57927	57927	57927
R2	0.007	0.085	0.227	0.082	0.222
R2 Adj.	0.006	0.043	0.190	0.040	0.184
Covariates			X		X
Mean F-stat 1st stage				3274	2994

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods.

Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

No effect on the probability of welfare or in-work benefits but an increasing amounts of social transfers To complete this picture, I run additional estimations of the dynamic intention-to-treat effects on additional outcomes that were pre-registered: i) Receiving RSA (Figure C.19), ii) receiving the in work benefit (PA) (Figure C.20) and iii) total social transfers (Figure C.21). I find no intention-to-treat effects of the programme on RSA or PA. At best, it suggests an increase of RSA take-up at the onset of the programme showing social workers' work of making sure all participants had access to all social transfers they were eligible to. This statistically insignificant *bump* can also be seen on the ITT estimates on the probability to report quarterly incomes (testing differential attrition) in Figure B.13.

However, this programme, which was expected to reduce reliance on social transfers, actually causes an increase in the amount received. Figure C.21 shows a steady trend in the intention-to-treat effect on total amount of cash transfers, which excludes 0 from the simultaneous confidence interval of the very last period. The table C.8 in the Appendix presents the aggregated effects over 6-months periods and finds a significant treatment effect on the treated of approximately € 100. Looking forward in time with the first three cohorts, Figure E.34 shows similar estimates as those with 4 cohorts up to 30 months, but the effects keep on increasing, excluding 0 from the simultaneous confidence intervals for several months.

Note that participants do not get more social transfers than when they were recruited. They lose less over time. Moreover, this variable does not account for changes in family size. Dividing total cash transfers by the contemporaneous number of consumption units, the mediated effect on cash transfers disappears entirely, as shown in Figure E.35. Thus, the lesser decrease in cash transfers compared with the counterfactual stems from changes in family size, as captured by the number of consumption units.

V.2 Slow-moving poverty rates and little distributional effects

I measure poverty as a threshold defined by the poverty line in 2019: € 1040 current. This value is actualised like the disposable income in $€_{2015}$ using the Consumer Price Index for the bottom income quintile.

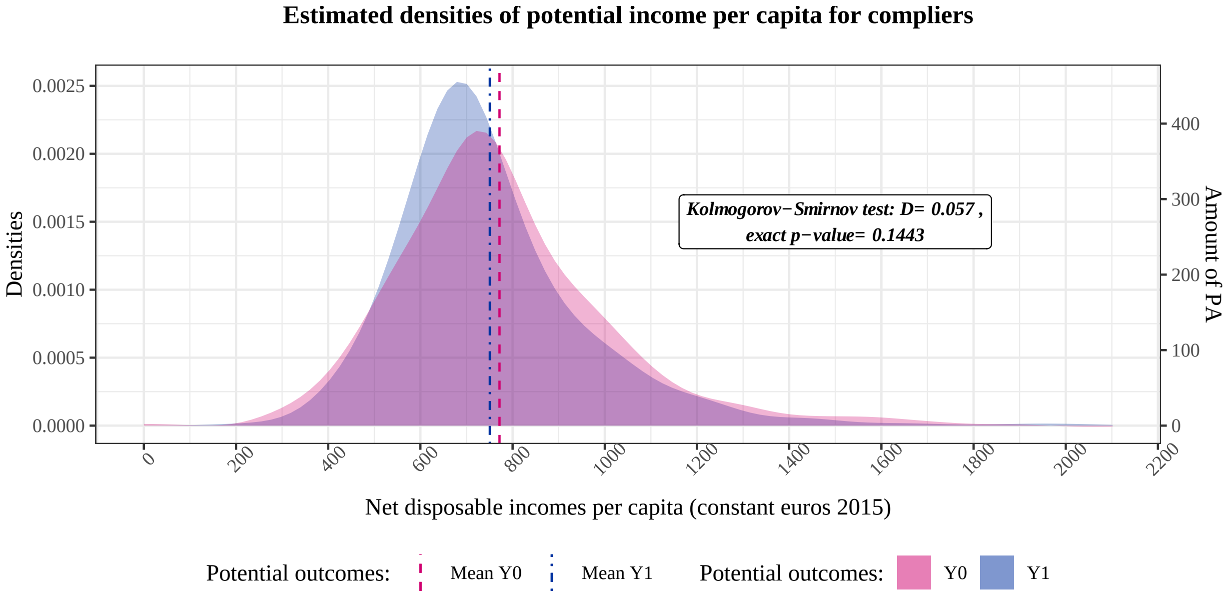
The programme slows the climb out of poverty Figure C.18 in Appendix C presents the average intention-to-treat of the risk of living with less than the poverty line. The top panel shows that at baseline, 98% of households live in poverty and, following the lock-in effect, we see the control group climbing out of poverty faster than the encouragement group in the first half of the programme, but a gap remains between the two groups that entirely close up at the very last period. The second panel estimates the intention-to-treat effects and shows a significant increase of poverty during the lock-in period. These estimates lack precision because of the number of parameters estimated. Aggregating the treatment effects over 6 months period and estimating the LATE, I find that the programme increases poverty incidence by 8 pp among compliers for the entire period of training, and exclude 0 from the adjusted 95% confidence interval (See Table C.9 in Appendix C).

Estimating participants potential incomes after the end of the programme To get a sense of the distributional effects of the programme, I estimate the counterfactual density of disposable incomes per consumption units from 18 to 30 months after the programme. To do that, I estimate equations (3b) and (3a) by 2SLS a 100 times on a grid of disposable incomes and using a kernel weighting of incomes as a dependent variable (See section IV for details).

The results on the years of observations after the end of the programme are presented in Figure 5. My previous estimations of the average effects on participants were small and insignificant. This figure shows that the potential distribution of incomes for treated and untreated compliers are quite similar over the entire distribution. A Kolmogorov-Smirnov test comparing the individual average outcomes over the period between the encouragement and control group cannot reject that the two distributions are drawn from the same distribution with a p-value of 0.144. However, potential outcomes of compliers when they are treated (blue density) are slightly more concentrated around € 700 while the counter-factual density is a bit thicker between € 800 and € 1100.

These results contradict my initial hypotheses. At the onset, my priors were that this programme *could* make a change and help some of these vulnerable families get out of poverty through employment. Policymakers spent four times as much as the usual budget for social support to fund a programme designed by practitioners, with

Figure 5: Distribution of the change in potential incomes per capita of compliers over 16 to 28 months (One year) after random assignment



Sources: ALLSTAT, restricted sample from 18 to 30 months and removing outliers (over 5 000 euros or the 99.7 percentile)
This figure plots marginal potential incomes per capita for compliers. Treated densities are estimated using 2SLS regressions of the interaction of a kernel density function and potential participation on the latter instrumented by encouragement.
Models only control for block x cohort x time to event dummies.
Untreated densities are estimated by replacing participation with one minus participation in this 2SLS procedure. All models use a Gaussian kernel and the Silverman (1986) rule of thumb bandwidth.
These estimates use a bandwidth of 64.3
The Kolmogorov−Smirnov test is run on a collapsed dataset of individual mean of the outcome over the period and compare the empirical CDF by encouragement group.
I test the sharp null that that incomes per capita are drawn from the same distribution using 99999 Monte carlo permutations.

a strong emphasis on *evidence based* social support. Yet, not a single assigned quantitative objective has been achieved. The programme seemed to have a positive impact on the first cohort but new data rejected this first initial findings. Analysis over a longer time frame for the first three cohorts show no improvement. However, the Covid-19 may have simply annihilate any benefit of the programme for the early cohorts. These conclusions will be confronted in future analysis including data from the fifth cohort and surveys.

It is worth noting that our results are in total contradiction with those of the qualitative evaluation. In the next section, I discuss plausible explanations, with a strong emphasis on endogenous selection.

VI Selection biais and heterogeneity

In this section, I investigate plausible mechanisms for the absence of effect on employment through two main points: self-selection and heterogeneity. First, I estimate difference in differences models comparing participants with control groups as if I had no randomised experiment. While the encouragement group has been randomly selected, participants did self-select into the treatment. The instrumental variable framework infer the share of compliers in the encouragement group and reweight the intention-to-treat effect by this proportion. If we had not randomise, we would not know the share of compliers and need an alternative identification strategy: typically matched difference-in-differences, In the economic literature, most researchers do not have the opportunity to run experiment of welfare to work programmes. There are far more quasi-experimental research typically using a difference-in-differences strategy to identify treatment effects. For instance, among the 39 studies included in the systematic review of ALMP for the unemployed of Filges et al. (2015), 25 used a timing-of-event and/or matching design while 14 were based on RCTs. If instead of comparing encouraged and control, I compare participants with another comparison group, the doubly-robust estimand of Callaway and Sant'Anna (2020a) could recover the average treatment effect on the treated under conditional parallel trend. This assumption is arguably much stronger than the instrumental variable framework which essentially rely on exclusion of alternative causal path between encouragement and outcomes but through participation. Here, a causal interpretation of these estimates requires that in the absence of treatment, participants would have had the same average outcomes than the comparison group, conditional on covariates. This assumption will be violated if there are unobserved time varying confounders. In contrast, estimations using the encouragement as an instrument and assuming exclusion hold true retrieve the average treatment effect on the treated. The comparison between the two estimates are therefore informative of the importance of selection bias.

VI.1 Self-selection: Observing the unobservables

To analyse the endogenous selection, I estimate difference-in-differences models of the effect of the programme on employment comparing participants with i) never-takers, ii) all non participants, iii) the control group, iv) the control group matching participants on the set of selected covariates. Never takers are those of the encouragement group who refused the programme and suffer from selection bias most. The control group contains never-takers and compliers - although I don't know who is who. The estimates with control units are biased by never-takers. Matched observations from the control group is the best I can do without randomisation.

The programme attracts those with highest employment potential The top panel of the Figure 6 presents the average employment by actual participation, splitting the encouragement group between compliers and never-takers. It corresponds to what Susan Athey and Imbens (2017) call *as treated* or *per protocol* analysis. It shows that participants have a much higher employment rate after the programme ends than the average of the control group and the never-takers. The employment rate of the latters is much lower than that of the average control group and we have seen earlier that the weighted average of treated and never-takers equals that of the control group.

This simple plot gives a first mechanism for the absence of effect: the programme attracted the single mothers who were most likely to find a job whether they participated or not. This selection effect is very strong and suggests that most single mothers who *would* have work enrolled in the programme. The qualitative evaluation entirely missed the selection bias in their analysis and instead, saw the effects of the programme in the large share of participants working. The complementarity of both approach is precious to understand what happens *under the hood* and provide clear messages for policy makers.

In the bottom panel, I estimate difference-in-differences using the modified CS (2020) estimand discussed above and use different control groups. Since I do not compare participants by block, I use cluster robust standard errors adjusted at the household level to account for serial correlation. First, I estimate the model *as if I had no control group* comparing treated and never-takers. This model is the most heavily biased by endogenous selection and yields the highest estimates. Then, adding the control group reduces the estimated difference to 10 percentage points. Third, removing never-takers from the comparison group takes out a large share of the observed differences. Finally, flexibility controlling for the chosen set of observables with the doubly-robust estimand of Callaway and Sant'Anna (2020a) further reduced the estimates and confidence interval no longer exclude 0.

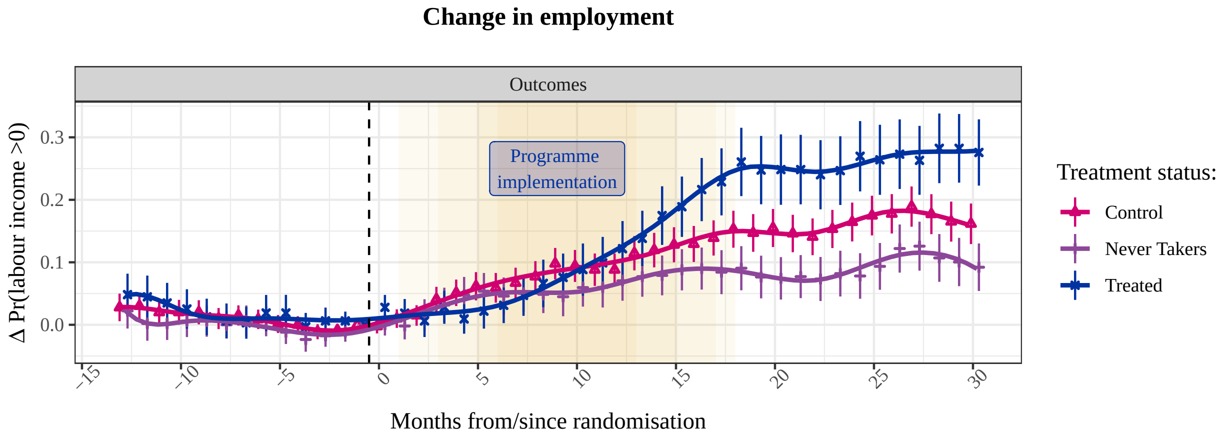
Visually, the event study is very convincing and the world is full of incentives to choose this model (Brodeur, Cook, and Heyes 2020). Pre-registration of the protocol, empirical strategy and main outcomes really made a difference. Over the past few years, it enabled me to dodge political pressures and critics on the choice set of outcomes. They also protected me from *HARKing*[48] my own research[49]. Difference-in-differences estimates are indeed very convincing. Estimates on lead months are close to 0 without trends, there is small lock-in and then a large increase in employment. However, all these results come from endogenous selection and collider bias due to the very nature of the target (Pearl 2009). The eligible population is conditioned upon receiving RSA for more than two years. Sampling is condition on past outcomes[50] and already partly controls for past employment spells. Moreover, the

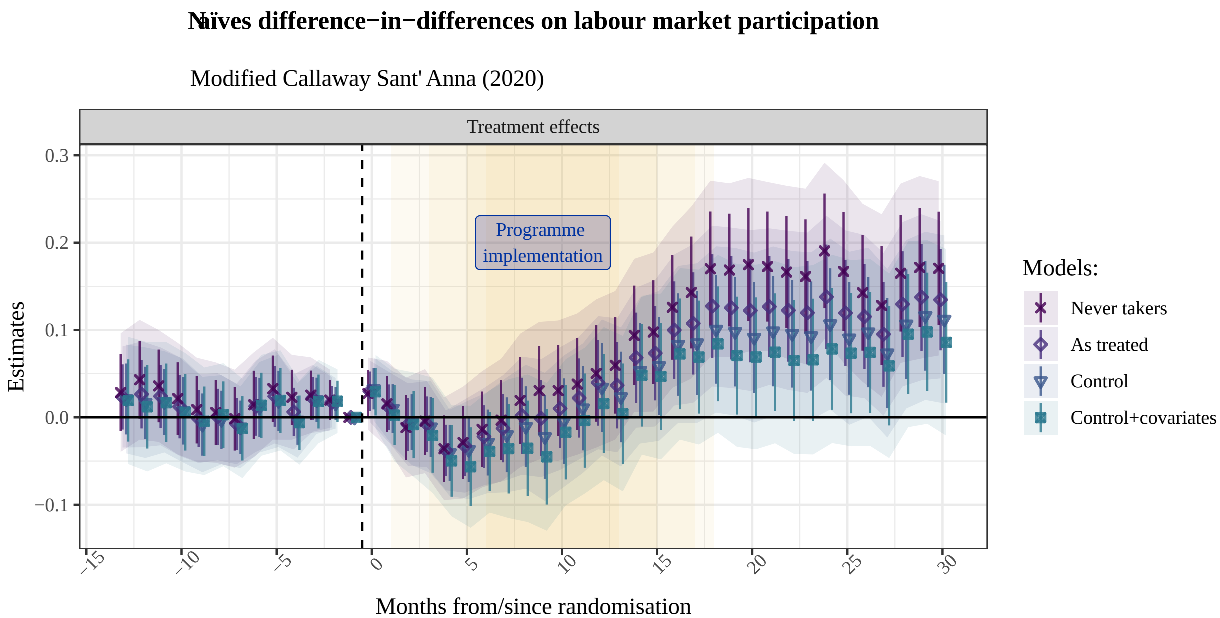
[48] Hypothesis After Results Are Known, see Kerr (1998) for a discussion.

[49] Not by lack of scientific integrity but because these results using fancy, flexible, doubly robust new methods where more in line with my priors on the effectiveness of the programme. Researchers are not exempt from confirmation bias and neophilia, especially when it comes with new tools in lieu of old-school *boring* regressions.

[50] Which is not a problem *per se*, it is the sub-population of interest.

Figure 6: Comparisons of labour market participation by treatment status and difference-in-differences.





Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the long difference between employment at any relative month and the month before randomisation.
Top panel:
− Points indicate simple means over cohorts 2018 to 2021 in relative time since randomisation with 95% error bars by treatment status.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
Separate difference−in−differences for each cohort using the doubly robust estimator proposed by Sant' Anna and Zhao (2020) aggregated following Callaway−Sant' Anna (2020), i.e. separate DiD for each cohort and weighted average of cohort x time treatment effects.
Covariates are measured at the month before random assignment and include baseline level, number of years receiving RSA, number of children and unemployment registration status (uninteracted blocking variables). French citizenship, High/Low education, favourable assessment, receiving each social transfers, child support, children between 3 to 5 and at least one child over 16, quartiles of age, income per capita, taxable income. The error bars indicate the 95% confidence intervals based on cluster−robust standard error adjusted at the household level. The shaded areas represent the FWER adjusted 95% confidence levels estimated using wild cluster bootstrap.

variables used to define blocks are not fixed attributes. For instance, the blocks of registered unemployed parents at the time of randomisation can be composed of long-term unemployed and very recent registrations. Without the experiment, this variable would be considered *endogenous* and require its own instrument. Said differently, eligibility varies with time and attributes measured at $t = c$ can be affected by past outcomes. In this setting, testing pre-trend can be very misleading. This is why I talk about *attributes* rather than characteristics: covariates are just observables at some points in time, held fixed and used to remove variation correlated with these attributes. Once we simply follow households, there are no other control and time varying differences between participants and others can affect the outcomes.

Difference-in-differences would reject the experimental results with 95% confidence. Table 3 displays the difference-in-differences estimates aggregated over months using long difference in employment as dependent variable, replacing the encouragement by participation in equation (7) and using only the control group for comparison (hence removing never-takers from the sample). The model with minimal covariates only includes specific trend by baseline outcome level. Models with covariate include the same set used in every estimations (see the figure notes). I compare these estimates with the ATT estimated by instrumental variables using long difference in employment as the outcomes. I do not control for baseline level in the first IV model to have no restriction on the treatment effect heterogeneity. Adding covariates does imply constant treatment effects but increase precision. I also present the estimate of the average *missing* compliers' potential outcomes $Y(0)$ following A. Abadie (2003). It is obtained using $(1-D_{ijc})Y_{ijcm}$ as dependent variable and $(1-D_{ijc}) \times S(m)$ instrumented by $\tilde{Z}_{ijc} \times S(m)$, the demeaned instrument for the relevant period. It shows that in the lock-in period, an average of 20% of participants would have worked had they not signed in the programme, although the average of the treated group is about 13%. Comparing estimations between instrumental variables and difference-in-differences, the estimated coefficients are generally of opposite sign and the 95% confidence intervals for the difference-in-differences estimates controlling for Family-wise error rates exclude the experimental point estimates from the lock-in period to the end. In other words, the best one could do with non-experimental methods cannot recover consistent estimates. Although not surprising (Lalonde 1986; J. J. Heckman, Lalonde, and Smith 1999; Card, Kluve, and Weber 2018), these results are worrisome if one thinks of the share of published research using such strategies to estimate the effect of similar programmes. I come back to this point in the discussion.

Table 3: Aggregated treatment effects on the treated on labour market participation

	Compliers' Y(0)	**TSLS**		**DID**	
		No covariates	*Covariates*	*minimal covariates*	*All covariates*
[-7 ; -1 [	0.099*** (0.031)	-0.024 (0.034)	-0.017 (0.023)	0.003 (0.010)	0.003 (0.010)
	[0.022, 0.175]	[-0.109, 0.061]	[-0.076, 0.042]	[-0.023, 0.029]	[-0.022, 0.029]
	adj.p.val. = 0.006	*adj.p.val. = 0.927*	*adj.p.val. = 0.944*	*adj.p.val. = 1.000*	*adj.p.val. = 0.999*
[0 ; 6 [	0.120*** (0.027)	-0.036 (0.030)	-0.039 (0.026)	-0.016 (0.013)	-0.018 (0.013)
	[0.053, 0.188]	[-0.113, 0.040]	[-0.106, 0.028]	[-0.051, 0.018]	[-0.052, 0.016]
	adj.p.val. = 0.000	*adj.p.val. = 0.627*	*adj.p.val. = 0.473*	*adj.p.val. = 0.693*	*adj.p.val. = 0.577*
[6 ; 12 [	0.208*** (0.036)	-0.073** (0.034)	-0.093*** (0.035)	-0.022 (0.016)	-0.025 (0.017)
	[0.117, 0.298]	[-0.159, 0.013]	[-0.184, -0.003]	[-0.064, 0.020]	[-0.069, 0.019]
	adj.p.val. = 0.000	*adj.p.val. = 0.128*	*adj.p.val. = 0.039*	*adj.p.val. = 0.607*	*adj.p.val. = 0.523*
[12 ; 18 [	0.261*** (0.045)	-0.016 (0.042)	-0.032 (0.042)	0.044* (0.023)	0.038* (0.022)
	[0.149, 0.374]	[-0.121, 0.090]	[-0.139, 0.074]	[-0.015, 0.102]	[-0.019, 0.096]
	adj.p.val. = 0.000	*adj.p.val. = 0.997*	*adj.p.val. = 0.927*	*adj.p.val. = 0.241*	*adj.p.val. = 0.355*
[18 ; 24 [	0.341*** (0.047)	-0.022 (0.047)	-0.052 (0.042)	0.076*** (0.025)	0.071*** (0.025)
	[0.224, 0.458]	[-0.140, 0.096]	[-0.160, 0.056]	[0.012, 0.140]	[0.007, 0.135]
	adj.p.val. = 0.000	*adj.p.val. = 0.989*	*adj.p.val. = 0.665*	*adj.p.val. = 0.012*	*adj.p.val. = 0.023*
[24 ; 30 [	0.365*** (0.047)	-0.021 (0.049)	-0.036 (0.048)	0.077*** (0.028)	0.072** (0.029)
	[0.247, 0.482]	[-0.146, 0.104]	[-0.160, 0.088]	[0.004, 0.151]	[-0.003, 0.147]
	adj.p.val. = 0.000	adj.p.val. = 0.994	adj.p.val. = 0.940	adj.p.val. = 0.034	adj.p.val. = 0.069
Num.Obs.	56749	56749	56749	39344	39344
R2	0.120	0.106	0.274	0.240	0.266
R2 Adj.	0.079	0.064	0.239	0.188	0.212
Baseline level			X	X	X
Covariates			X		X
Mean F-stat 1st stage	3198	3223	2928		

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: the dependent variable equals 1 when the parent has positive labour incomes, 0 otherwise. Compliers' average obtained by TSLS of DY on D instrumented by centered instrument. Second and third columns report TSLS estimates of the ATTs with or without adding covariates in the model. I report the average of the F-stats for the first stages of all treatment periods. Columns DID remove nevertakers from the sample and compare treated and controls in difference-in-differences. Covariates are measured at the month before randomisation and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. Observations that are treated in the cohort 2022 are dropped once they are resampled.

VI.2 Puzzling heterogeneous treatment effects by children at baseline

In Appendix D, I provide estimates of the average treatment effect on the treated for employment and disposable income per consumption unit, categorized by various subgroups determined by blocking variables. Two additional exploratory analyses, not pre-registered, are also presented. All models incorporate covariates to enhance precision and are estimated over 6-month periods using Two-Stage Least Squares (TSLS) with demeaned instruments and block fixed effects.

Heterogeneous effects by cohort Figure D.22 illustrates the heterogeneous effects by cohort. Notably, the 2018 cohort exhibits negligible impacts on both employment and incomes throughout the entire period. Conversely, the 2019 cohort experiences a pronounced lock-in effect on employment, driven by a temporarily higher job finding rate for the control group post-PA reform. This effect diminishes by the end of the training, with no significant impact on disposable incomes. The 2020 cohort shows minimal effects on employment but a significant lock-in effect on incomes, dissipating rapidly. In contrast, the 2021 cohort reveals a distinct pattern, with negative employment effects intensifying post-lock-in, implying a -15 pp lower probability of employment 2 years after randomisation. While no employment effects are observed in the last 6 months, disposable incomes follow a similar trajectory, with a significant increase over the initial 18 months. There are no significant effect or heterogeneity on disposable income after 2 years since random assignment.

Treatment effects by number of children at baseline Figure D.23 presents the ATTs by the number of children at baseline. Participants with one child experience a significant, lasting negative effect on employment, peaking at about 17 pp by the programme's end. This group also encounters a substantial lock-in effect on disposable incomes, approximately €150. However, the income effect dissipates post-programme, with no lasting impact. Conversely, those with three or more children exhibit the opposite pattern, though estimates lack significance. They experience no lock-in effect on employment or income, with a positive employment effect stabilising around 10 pp higher (CI between -5 and 35 pp). The effect on disposable incomes is increasingly negative but not statistically significant.

Treatment effects by other subgroups Comparisons by the number of years receiving RSA and registration with the Employment agnecy (Figure D.24 and D.26 respectively) reveal minor differences, with slightly worse effects for those registered at the Employment agnecy and those with fewer than 5 years receiving RSA.

Figure D.25 illustrates outcomes by social workers' initial assessments, revealing more negative effects among those with a favourable assessment, while others experience increasingly negative effects on income per consumption unit with little to no impact on employment.

Finally, estimations by income levels at baseline presented in Figure D.27 indicate substantial negative and lasting effects on employment for those with incomes below the median. Interestingly, this group, more likely to participate, sees no effect on incomes. In contrast, those with higher baseline incomes exhibit no employment effects but face a significant decrease in disposable incomes by about €100 per month.

In summary, while the programme demonstrates little to no average effect on employment and disposable incomes, the heterogeneous analyses unveil noteworthy patterns, particularly regarding the number of children at baseline. Altogether, these results suggest heterogeneous changes in the composition of incomes which constitute a new puzzle that I address in Galitzine and Heim (2024).

VII Discussions

In this paper, I analysed the effects on labour market participation and poverty of an intensive welfare-to-work programme for single parents on long-term welfare in France. This experiment relied on a staggered block-randomised encouragement design to recruit participants from 2018 to 2022, building a dataset of 2073 households from 5 cohorts. I follow the 4 first ones (1671 households) with panel data from administrative records up to 30 months after random assignment. The main findings can be summed-up as followed:

1) The take-up is 38% on average, increasing from 28% to 48% from the firth to fourth cohort. This improvement most likely stems from adjustments to the recruiting process including i) more threatening invitations ii) information sessions moved to the programme's premises iii) change from collective to individual face-to-face meetings with project managers and iv) testimonies from former participants.
2) Compliers are more likely to be in their thirties, among the poorest, have less than a high school diploma, and be registered to the Employment agency.
3) The programme slows the job finding rate during its first half, causing a strong lock-in effect showing up on poverty rate, disposable income and employment. These anticipated negative effects fade-out by the end of the programme and there are no average effect in the post-treatment period.
4) The programme gradually increases the amount of cash transfers received, although this effect is entirely mediated by increasing family size.
5) Analysis of the heterogeneity of treatment effects reveals puzzling patterns, especially among high/low income and number of children at baseline, where effects on disposable incomes and labour market participation have diverging patterns, suggesting changes in the composition of incomes, corroborated by the increase in cash transfers at the end of the period.
6) The programme attracts those with highest potential employment levels, but do not increase labour market participation. The selection bias is so strong that estimates using the next-best identification strategy implemented with modern robust estimands fail to include the experimental results in the confidence intervals. Without random assignment, one would have wrongly concludes that the programme increases employment.

The effects of this programme on labour market participation and disposable income of poor single mothers in France are very negative. It failed at increasing employment or get these single parents out of poverty. In this section, I discuss plausible interpretations, threats to their validity and some policy implications.

A diversity of profiles with heterogeneous reactions In this experiment, compliers seem to be divided into two groups: the poorest with low education on the one hand, and those closer to the labour market on the other hand. For the formers, the programme reduces employment but has no effect on disposable incomes while for the latter, it has no effect on employment but significantly reduces their incomes. This can only mean that the poorest found other sources of incomes while those closest to the labour market earn less than they would have had they not participated. The total transfers by Family allowance funds are higher due to higher number of consumption units. This means that the programme affected family size and composition, possibly increasing cohabitation, fertility or through lower custody loss. Understanding these complex reactions are out of the scope of this study. In Galitzine and Heim (2024), we analyse these mechanisms and link compliers reactions to the incentives of the tax-benefit system, which, as it turns out, has a 10-15% implicit marginal tax rate at part-time minimum wage, and a 70-75% at the full time minimum wage. Estimating potential labour incomes distribution, we show that treated compliers bunch at part-time minimum wage and find a hole in the distribution for incomes higher than minimum wage. The qualitative evaluation also reports that among participants of the 2019 cohort on permanent or fixed-term contracts, 72% work part-time (58% on fixed-term part-time contracts) and 75% say they choose to do so[51]. Moreover, incentives are different by number of children and cohabitation with a partner, and we also find consistent treatment effects on the probability of re-partnering and de-cohabitation with older children. However, we confirm the absence of treatment effects on disposable incomes over the entire distribution, no matter how many children they had at baseline. Together, these two studies show that the programme helped single mothers optimise their circumstances, reducing work hours

[51] Source from the internal data from the Departmental council. 5% say they they did not choose it and 20% did not answer.

and incomes for those who would have worked, while also adjusting family size and composition, such that the resulting higher social transfers compensate lower labour incomes, maintaining the same standards of living as in the counterfactual. This may also include income effects when social workers helped participants access local aids or discounts - things I don't observe in the data.

Mothers and children in poverty and job quality Although the programme helped participants optimise, 89% remains in poverty 30 months after random assignment and 66 % have no job. We supported 328 single parents with their 672 children and 30 months after random assignment, 277 participants and their 568 children remain poor. Even among those who work, poverty remains strikingly high with 75% working poors among participants 30 months after randomisation. Despite social transfers, low-paid jobs do not provide sufficient income to support families and often do not meet parents' constraints (P. M. Evans 2007 ; Jaehrling, Kalina, and Mesaros 2015; Millar et al. 2018; Van Winkle and Struffolino 2018). Moreover, they do not allow mothers to build human capital (Blundell et al. 2016) and prevent them from investing in their children' education (Løken, Lommerud, and Holm Reiso 2018). Moreover, exposure to poverty during childhood significantly impacts various aspects of psychological well-being trajectories and learning abilities, with long term consequences (Lucas et al. 2008; Duncan, Ziol-Guest, and Kalil 2010; Duncan et al. 2012; G. W. Evans and De France 2022). Conversely, activation with high social transfers has been shown to durably get single parents out of poverty (Markussen and Røed 2016).

In the qualitative study, some participants felt that their aspirations were not sufficiently considered. This includes mothers who were directed towards jobs or sectors that did not interest them or felt "pushed" into employment despite their difficulties. In fact, a large share of offers were jobs in the personal care sector (home help and care assistants), in supermarkets (checkouts or stock replenishment), or in cleaning services. Some individuals were not ready or equipped to face the working conditions proposed in low-skilled sectors and/or the time frame set by Reliance. In some cases, the person's initial project may have been out of sync with their current skills and/or the functioning of the job market, leading to a redirection towards job offers deemed more "realistic" by social workers. This paternalistic view is well documented and tends to create a sense of loss of control in people's life. Campbell et al. (2016) reviewed qualitative studies on welfare-to-work programmes for single mothers, finding that programmes' demands often clash with parenting duties, resulting in precarious and low-paying jobs. Health issues like stress and depression are common, but some participants experience increased self-esteem. However, these programmes can reduce control over employment and childcare, contributing to health concerns. Similar findings were noted in the literature review of Baronnet et al. (2021).

Unresolved issues and possible hidden effects For others, Reliance's support helped realise that employment is not their short or medium-term goal. In most situations, participants are still overwhelmed by issues that strongly impact their daily lives, preventing them from envisioning a return to employment: lack of personal housing, child placement, serious illness, a child's disability, etc. The persistence of these trajectories reveals the lack of insurance against the consequences of certain life shocks. These situations, described by Perrin-Heredia (2009), highlight that for the most deprived social categories, "*life accidents are not random. [They] are both more likely to experience these accidents […] and have fewer means to face them.*".

However, they may still benefit from the quality of social support, but not through employment. Indeed, the qualitative evaluation reports high satisfaction, increased motivation and self esteem, while reducing loneliness and isolation. PTSD and depressions were common among participants, some resulting from intimate partner violence and coercion (IPVC). Other research showed that such programme can have positive effects on IPVC-related PTSD (Meisel, Chandler, and Rienzi 2003; Perez, Johnson, and Wright 2012; Latzman et al. 2019).

The development of models to measure the effects of welfare programmes began in the late 1960s and early 1970s with the static labour supply model and was well worked out by the 1980s, extending to education, marriage, fertility savings and so on. And yet, as Chan and Moffitt (2018) point out, "*there has been far too little work on the dynamic aspects of labour supply choices in the presence of different kinds of programs (traditional welfare versus earnings subsidies, for example) where human capital, family structure, migration, occupational choice, and other lifecycle decisions are important*". Economics may well have missed important phenomena dismissing these dimensions. At this point, the lack of measures on other outcomes may put too much weight on the negative results of this evaluation. The analysis of surveys left for future work will help gauge the effects of the programme on subjective health and well-being.

The recruiting process was successful Targeting vulnerable populations often produces low take-up, which is attributed to various factors such as low monetary or utility gains, stigma associated with social programmes, monetary, non-monetary and opportunity costs of participation, imperfect information, administrative barriers, and inadequate eligibility measurements (Friedrichsen, König, and Schmacker 2018 ; Ko and Moffitt 2022). In this experiment, the lowest take-up is 8 points higher than a similar experiment from 2006 in Seine-Saint-Denis, the poorest suburb close to Paris (Crepon et al. 2013). After four years of implementation, we managed to enroll almost half of the 2021 cohort. For that, we played on both *threats* and *pull* levers in the recruiting process. First, we added an ambiguous yet threatening sentence on *"rights and duties"* in the invitation letter: a *threat effect* to foster participation. Second, we made recruiting sessions more warm, welcoming and individualised. We moved from collective information sessions in the *welfare-to-work* division of the Departmental council to individual face-to-face interviews with project managers in the newly renovated and well equipped premises of the programme. Additionally, former participants were involved in the recruiting sessions to answer questions and provide (positive) testimonies as peers. Higher take-up likely stems from these adjustments as corroborated by the qualitative evaluation. However, the Covid-19 pandemic and resulting restrictions of social contacts may have played a role as well, increasing the need to break isolation and loneliness. Furthermore, random sampling and assignment are very different from usual referral practices and may have reached households who would not have been proposed such programme.

Researchers and policy-makers have implemented various *carrots vs. sticks* schemes to foster participation in ALMP. Recent experimental works show that receiving information is not enough ; transaction costs and administrative barriers matters and human contacts help foster participation. For instance, Chareyron, Gray, and L'Horty (2018) evaluate how changing the content of the official letter sent to RSA recipients could foster participation in welfare-to-work programmes. Randomising wordings of different letter templates, they find no effect of either the "simplified letter" or the one highlighting the potential benefits of the programme. Two recent works analyse randomised interventions designed to increase take-up of the Supplemental Nutrition Assistance Program (SNAP) in the US (Finkelstein and Notowidigdo 2019) and all social benefits in France (Castell et al. 2022). Both use a variation of an information-only $vs.$ human assistance compared with a control group. Both find that human assistance increase take-up. While Finkelstein and Notowidigdo (2019) find that providing targeted information also increases take-up, Castell et al. (2022) find no such effect in their setting. Finkelstein and Notowidigdo (2019) further show that the compliers are initially better off than the control group, thus leaving out those who need it most or are more likely to benefit from it.

On the importance of high quality research designs This experiment was designed and analysed with great care. I used block-randomisation to ensure balance between groups, gain precision and be able to test the presence of heterogeneity. I pre-registered the design, main outcomes and estimation methods on Social Science Registry before having access to the data. The benefit of using a randomised experiment are particularly salient in this work. Without random assignment, the next best alternative to estimate the effect of the programme on participants yields estimates of the opposite sign excluding the experimental estimates from the confidence interval. This result is worrisome considering the lack of experimental evidence on such policies, the vulnerability of the target audience and the public spending. For instance, the review of Bergemann and Van Den Berg (2008) on ALMPs for women in Europe conclude that they are very effective but only 4 out of 39 included research use a randomised experiment, two of which show negative results. The evolution of the literature on modern difference-in-differences raises additional doubts on the reliability of these estimates (Roth et al. 2023).

Opportunity costs of public funds In this experiment, policymakers revealed their willingness to get to lift these families out of poverty, spending roughly € 2800 per participants, setting an employment target of 10 pp for the success. They were ready to support 10 single parents so that one would get a job on average, implicitly accepting a cost of € 28 000 per expected job. This exceeds the total employer cost of one year full-time job at the minimum wage (which is approximately €22,000 per year) while relying on highly uncertain effects. And in fact, the programme did not increase employment and increased social transfers instead. This echoes one of the arguments motivating the "*Territoire Zéro Chômeur*" project, where long-term unemployed are offered permanent contracts with full time minimum wage. This high stake experiment was not randomised but the statistical service of the Ministry of Labor (DARES) used matching to evaluate the effects and find significant positive effects on participants' employment and well-being (DARES 2021). Conversely, Kasy and Lehner (2023) evaluate a similar policy in Germany and use a very high standard research design. They look at the individual average treatment effect using a randomised-match-pair design, evaluate the aggregate effect using a pre-registered synthetic control at the municipality level and comparison to individual in control municipalities. The latter allows to identify spillovers. They find positive impacts of program participation on economic and non-economic wellbeing, but not on physical health or preferences. At the municipality level, they find a large reduction of long-term unemployment, and no negative employment spillovers.

It seems important to highlight the risk taken in pursuing activation policies and to question the motivations and justifications for these choices, considering the costs involved, the risk of failure, and the little consideration for high quality design. Could different use of this same budget reach policymakers objectives ? Could one do better ? For instance, several recent studies support the idea that a significant monetary transfer produces lasting positive effects on poverty (Jones and Marinescu 2022) and children's education (Barr, Eggleston, and Smith 2022) in very different contexts. Most recently, a large scale unconditional cash transfers randomised experiment in the US showed that it increases mothers spendings for and time spend with their children (Gennetian et al. 2022). However, these amounts must be sufficient to meet the needs of families, otherwise, such interventions have no lasting effect (Jaroszewicz et al. 2022). Overall, social assistance programmes in high-income countries are insufficient in preserving the health and well-being of socio-economically disadvantaged populations, indicating that the scope and generosity of existing programmes fall short in compensating for the negative consequences associated with poverty (Shahidi et al. 2019).

'*From welfare-to-work – and worries*' While the previous presidency of Emmanuel Macron has been analysed as a continuation of the *flexicurity à la Française* started by Nicolas Sarkozy (Gazier 2019; Askenazy 2022), his second term is strongly influenced by the *"Welfare chauvinism"* pressure of Radical right populists (Rinaldi and Bekker 2021). In the European parliament, these new movements have changed the content of debates on social policies. They adopt a very different position on Welfare States from traditional Right wing parties in that they use it as mean to redefine "*the people*" - "*insiders*" and "*outsiders*", "*deserving*" and "*underserving*" - through rights and obligations.

The new welfare reform is set to make social support *mandatory* and put the Employment agency in charge of it, with large sanctions. In light of these results, this policy is not only regressive, it goes against a large and growing body of evidences[52] to which this research add yet another null effect. In particular, the analyse of the effects of sanctions for RSA recipients by Chareyron, Le Gall, and L'Horty (2022) find increased registration to Employment agency and social support but also increased non-take-up of RSA. It also dismiss alternative use of public funds that showed promising results. For instance, recent randomised experiments in Europe indicate that removing *workfare obligations* can improve labour market outcomes (Verlaat et al. 2021), or be as effective as intensive monitoring while enhancing trust in public institutions (Betkó 2023), or that job guarantee improve individual welfare, reduce long-term unemployment without negative spillovers (Kasy and Lehner 2023).

A quarter century after "Anti-poverty for families the next century: From welfare-to-work – and worries" of Ellwood (2000) in the *Journal of Economic Perspectives*, it is worrying, indeed, that question for the next 25 years turned into: *Welfare-to-what ?*

[52] Control and sanctions may expedite return to employment but at the cost of reduced wage and job quality (Arni, Lalive, and Van Ours 2013). Existing literature suggests that harsher control can direct poor individuals toward disability pensions or leave them without income (McCRATE and SMITH 1998; Dwyer et al. 2020; Morescalchi and Paruolo 2020; de Gendre, Schurer, and Zhang 2022).

Appendix

A The Reliance experiment

A.I Recruiting participants

A) Invitation letter and invitation leaflet for the first cohort

Figure A.7: Model letter for recruiting the first cohort

Laxou, le 15 Février 2018

MODELE

Dossier suivi par
ARELIA – Dispositif RELIANCE
Tél : [redacted]

Tout sur le RSA en Meurthe et Moselle :
www.insertion.meurthe-et-moselle.fr

Objet : RSA – Information collective RELIANCE

Madame,

Vous êtes bénéficiaire du Revenu de Solidarité Active (RSA).

La Caisse d'Allocations Familiales, le Conseil Départemental de Meurthe-et-Moselle ainsi que la Caisse des Dépôts et Consignation mettent en place une **action d'accompagnement destinée aux chefs de familles monoparentales**. Dans ce cadre, vous avez été identifié(e) pour y participer.

Vous êtes invité(e) à une réunion d'information collective présentant cette action intitulée RELIANCE dont l'objectif est de favoriser, à terme, votre accès à un emploi ou une formation. Au cours de cette rencontre, un temps vous sera réservé afin d'évoquer votre situation, vos projets et vos modalités d'organisation.

Celle-ci aura lieu le :

Date : **Lundi 19 Mars 2018 à 09H30 à 11H00**

Lieu : **ARELIA RELIANCE**
9-11 rue Robert Schuman, 3ème étage
54500 VANDOEUVRE LES NANCY

(Voir plan au dos)

Merci de vous organiser pour vous rendre disponible. Néanmoins, ***en cas de difficultés de garde****, nous pouvons vous accueillir en présence de vos enfants.*

Si vous ne pouvez pas venir à cette réunion, nous vous demandons de téléphoner au [redacted] **réception de cette lettre**, pour nous en informer.

Dans l'attente de vous rencontrer, je vous prie de recevoir, Madame, l'expression de nos salutations distinguées.

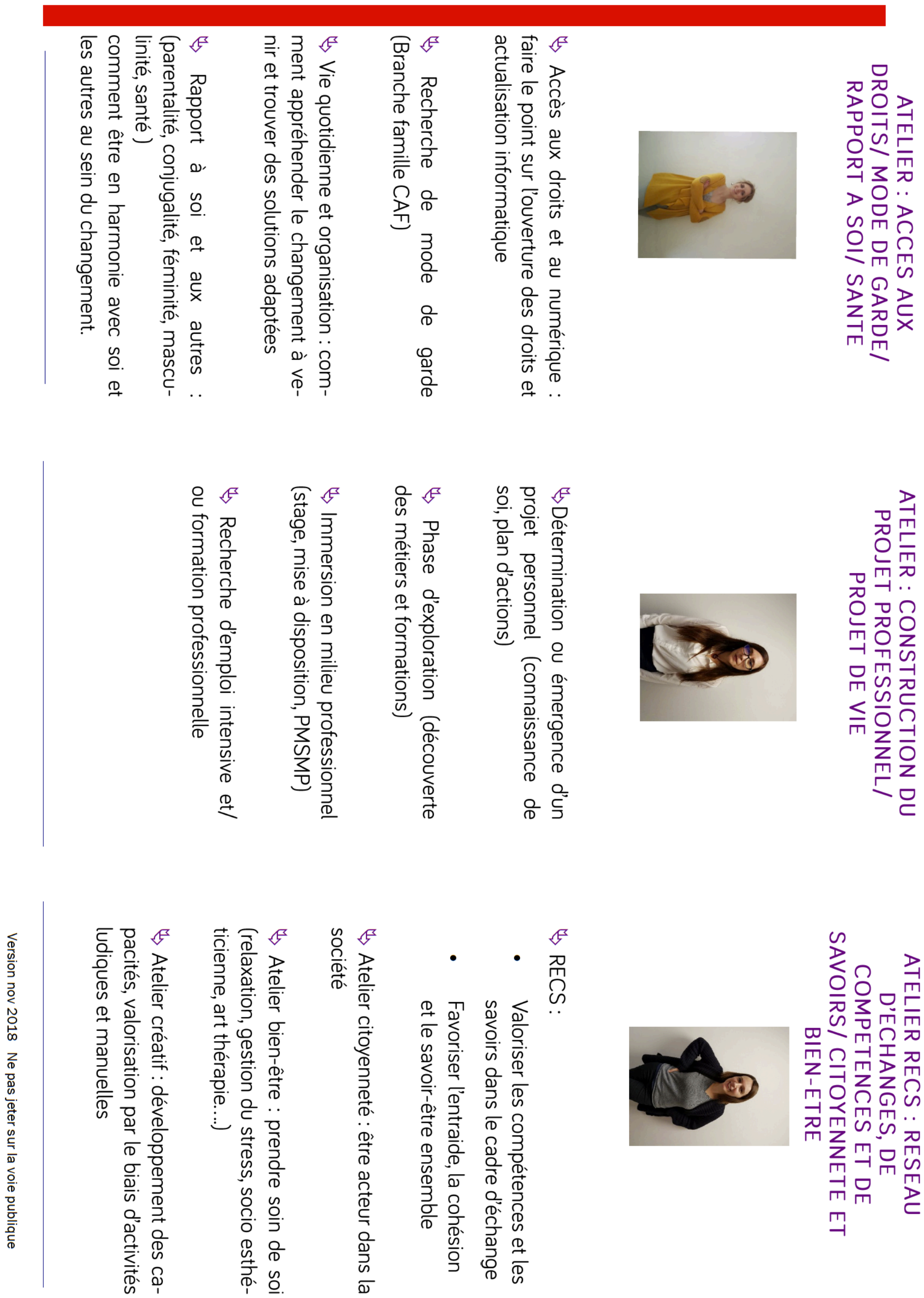
Pour le Président du Conseil Départemental et par délégation,

[redacted]
Responsable du Service Economie Solidaire et Insertion

Reliance est un dispositif porté par l'association Arélia, en partenariat avec Ulis et Ecoval

Figure A.8: Presentation leaflet 1/2

Figure A.9: Presentation leaflet 2/2



ATELIER : ACCES AUX DROITS/ MODE DE GARDE/ RAPPORT A SOI/ SANTE

- Accès aux droits et au numérique : faire le point sur l'ouverture des droits et actualisation informatique
- Recherche de mode de garde (Branche famille CAF)
- Vie quotidienne et organisation : comment appréhender le changement à venir et trouver des solutions adaptées
- Rapport à soi et aux autres : (parentalité, conjugalité, féminité, masculinité, santé) comment être en harmonie avec soi et les autres au sein du changement.

ATELIER : CONSTRUCTION DU PROJET PROFESSIONNEL/ PROJET DE VIE

- Détermination ou émergence d'un projet personnel (connaissance de soi, plan d'actions)
- Phase d'exploration (découverte des métiers et formations)
- Immersion en milieu professionnel (stage, mise à disposition, PMSMP)
- Recherche d'emploi intensive et/ ou formation professionnelle

ATELIER RECS : RESEAU D'ECHANGES, DE COMPETENCES ET DE SAVOIRS/ CITOYENNETE ET BIEN-ETRE

- RECS :
  - Valoriser les compétences et les savoirs dans le cadre d'échange
  - Favoriser l'entraide, la cohésion et le savoir-être ensemble
- Atelier citoyenneté : être acteur dans la société
- Atelier bien-être : prendre soin de soi (relaxation, gestion du stress, socio esthéticienne, art thérapie...)
- Atelier créatif : développement des capacités, valorisation par le biais d'activités ludiques et manuelles

Version nov 2018 Ne pas jeter sur la voie publique

B) Invitation letter for the second cohort

Figure A.10: Letter in the 2019 and subsequent cohorts

Laxou, le 09 Novembre 2018

Madame XXXXXXXX XXXX
XXXXXXX
XXXXXXX

54000 NANCY

Dossier suivi par
ARELIA – Dispositif RELIANCE
Tél : [redacted]

Tout sur le RSA en Meurthe et Moselle :
www.insertion.meurthe-et-moselle.fr

Objet : RSA – Information collective RELIANCE

«Civ»,

La Caisse d'Allocations Familiales, le Conseil Départemental de Meurthe-et-Moselle ainsi que la Caisse des Dépôts et Consignation organisent une **action d'accompagnement destinée aux chefs de familles monoparentales** appelée RELIANCE. Dans ce cadre, vous avez été identifié(e) pour y participer.

L'objectif de cette action d'insertion est de vous soutenir dans la construction et la réalisation d'un projet professionnel personnalisé. Elle comprend également de nombreux ateliers organisés autour de l'accès aux droits, à la santé, à la citoyenneté, à la garde d'enfants…

Une information détaillée du dispositif vous sera présenté le :

Date : **Lundi 10 Décembre 2018 de 09H30 à 11H00**

Lieu : **ARELIA RELIANCE**
9-11 rue Robert Schuman, 3ème étage
54500 VANDOEUVRE LES NANCY

(Voir plan au dos)

Je vous informe que **votre participation à cette réunion s'inscrit dans vos obligations relatives à la législation du dispositif RSA.**

Merci de vous organiser pour vous rendre disponible. Néanmoins, **en cas de difficultés de garde**, nous pouvons vous accueillir en présence de vos enfants.

Si vous ne pouvez pas venir à cette réunion, nous vous demandons de téléphoner au [redacted] **réception de cette lettre**, pour nous en informer.

Dans l'attente de vous rencontrer, je vous prie de recevoir, Madame, l'expression de nos salutations distinguées.

Pour le Président du Conseil Départemental et par délégation,

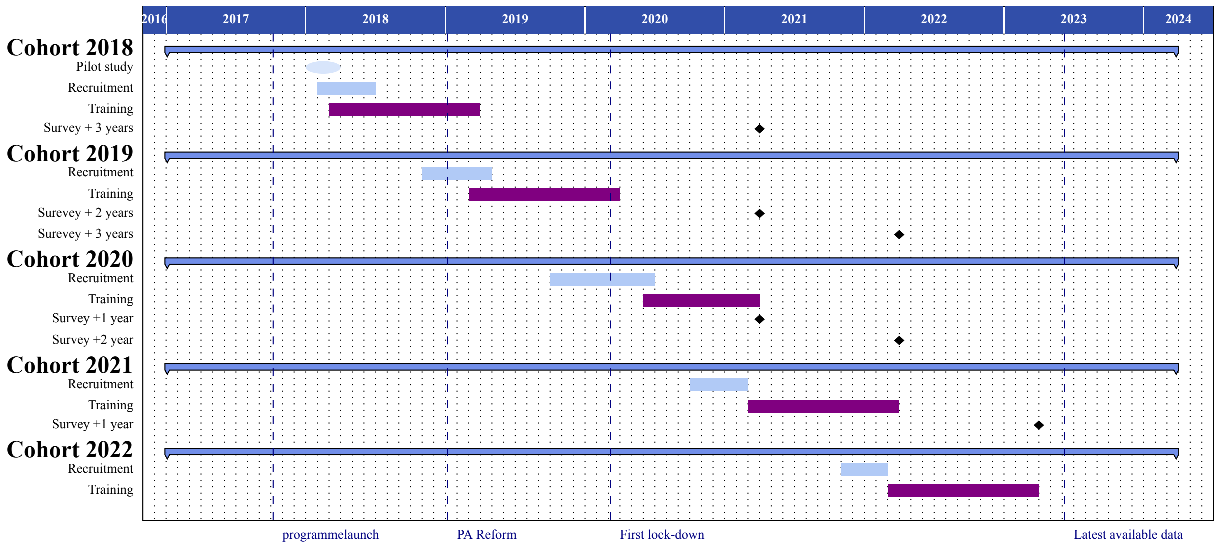
[redacted]
Responsable du Service Economie Solidaire et Insertion

Reliance est un dispositif porté par l'association Arélia, en partenariat avec Ulis et Ecoval

A.II Implementation details

A) Timeline of the experiment

Figure A.11: Timeline of the experiment



B) Adaptation of the programme during the pandemic

The Covid-19 epidemic and the sanitary restrictions implemented in France from March 2020 onwards had a significant impact on the course of support and, more importantly, on the living conditions, prospects, and opportunities for this particularly vulnerable population.

The year 2020 was marked by the impact of lockdown measures, curfews, and activity restrictions in certain sectors, to varying extents. The exceptional measures taken by authorities to support businesses and households likely mitigated the recessive effects of the health crisis to some extent. However, the economy has been durably affected, and vulnerable populations have been particularly impacted (Duvoux and Lelièvre 2021).

The first lockdown was implemented just as the participants of the 2020 cohort were starting their support journey. One challenge for the Reliance team was to maintain a connection despite the lockdown. To achieve this, a coordinator was assigned to regularly call all participants and answer their inquiries during extended time slots, including evenings and weekends. Social media, particularly a Facebook group, was utilised, with the coordinator regularly sharing diverse content ranging from recipes and activities to job-related quizzes, documentation, and information, etc.

The Reliance team also identified limited access to digital resources for some participants. Consequently, computer equipment (laptops, tablets) was provided for free to certain participants with the support of the Caf and the Departmental Council. Social workers also arranged for food parcels to be delivered to homes and printed important documents for beneficiaries who no longer had access to printing facilities. A mask manufacturing plant offered employment opportunities for participants to make masks, and volunteer sewing workshops for mask-making were organised.

The qualitative evaluation highlighted that contact was generally well maintained, and those interviewed felt supported. However, activities initiated before the lockdown were put on hold (such as driver's licence exams, workplace internships, access to rights, etc.), and the Reliance team noticed a greater level of "disengagement," particularly less participation in collective activities despite the implemented safety measures. FORS (2020) notes explanations from participants, such as fear of rejoining groups in the current health context or overwhelming anxiety. Overall, the lockdown period was anxiety-inducing for Reliance's supported individuals, who once again found themselves isolated, potentially exacerbating their difficulties. However, FORS also highlights that this period helped re-motivate some participants despite the context.

B Data and presentation of the variables used in the analysis

B.I Data sources

Data quality and attrition I observed all households that received at least one CAF payment in the year, but some data are collected more frequently than others and some come from other administrations (*e.g.* taxes or unemployment benefits). Those who apply for RSA, PA, or the disability allowance for adults (AAH) must report their incomes and situation every quarters. Most variables come from these quarterly reports after statistical consolidations[53]. Families in our sample have been registered for RSA for at least two years and are, therefore, used to fill these quarterly declarations. On average we have complete data for 92% the period after randomisation. Only 4.1% files are lost. Figure B.12 compares the attrition rate across cohorts and samples and shows that the missing income or missing files are very similar for the encouraged and control groups, whereas those excluded by social workers exhibit higher attrition patterns. In figure B.13, I estimate the effect of encouragement on the probability of observing incomes each month and show the event study-plot[54]. Overall, there are no differential attrition between the treatment arms although there is a small bump of higher reporting at the early stage of the programme.

Sample restrictions The block sizes vary a lot and some of them have four observations or less. There are few blocks with less than 2 observations in each treatment arms. In order to compute cluster-robust standard errors at the block level, I drop observations from these blocks from the analysis. This deletion is innocuous and reduce the full database from 103066 individual $\times$ month observations to 102749. Only 1 treated household is dropped.

In addition, I restrict the window of analysis to the timeframe with overlapping relative times since randomisation to avoid composition effects. I can only observe the 2018 cohort from -13 month before randomisation. For the 2021 cohort, data go up to 30 months after randomisation. Estimations over 4 cohorts are always over this interval. I also look at the effects over the first three cohorts over a longer period.

Finally, the 2022 cohort includes 127 households from the control groups of the previous cohorts. For instance, in the 2018 cohort, 24 were re-sampled, among which 12 were invited with the fourth cohort, 5 of which enrolled. In the main analysis, the observations of the encouragement group among those re-sampled are dropped, creating a slight compositional change in the analysis. There data are analysed with the 2022 cohort. Those resampled but remaining in the control group are duplicated in the dataset, once in their original cohort and once in the control group of the 2022 cohort.

[53] For each month, there are three files corresponding to three extraction delays: FR1, FR2, and FR6, corresponding to 1, 2, and 6 months of recall. The quality of information about beneficiaries and their rights to benefits increases with extraction delay and these are certified data The later the files, the better they include "latecomers". "Latecomers" occur because parents did not send all information on time or because there were administrative controls, delays in mail treatment, errors and other procedures.

[54] This model is estimated exactly like other models for intention to treat analysis. See section IV for more details.

Non-response: casewise deletion These administrative data have little problem of missing variables apart from attrition which is independent from the treatment. There are however some that have incoherent values and missing observations coded 99999 in original files when, for instance, parents stop for a quarter to send their income. So far, I simply delete missing cases.

B.II Description of the main variables of interest

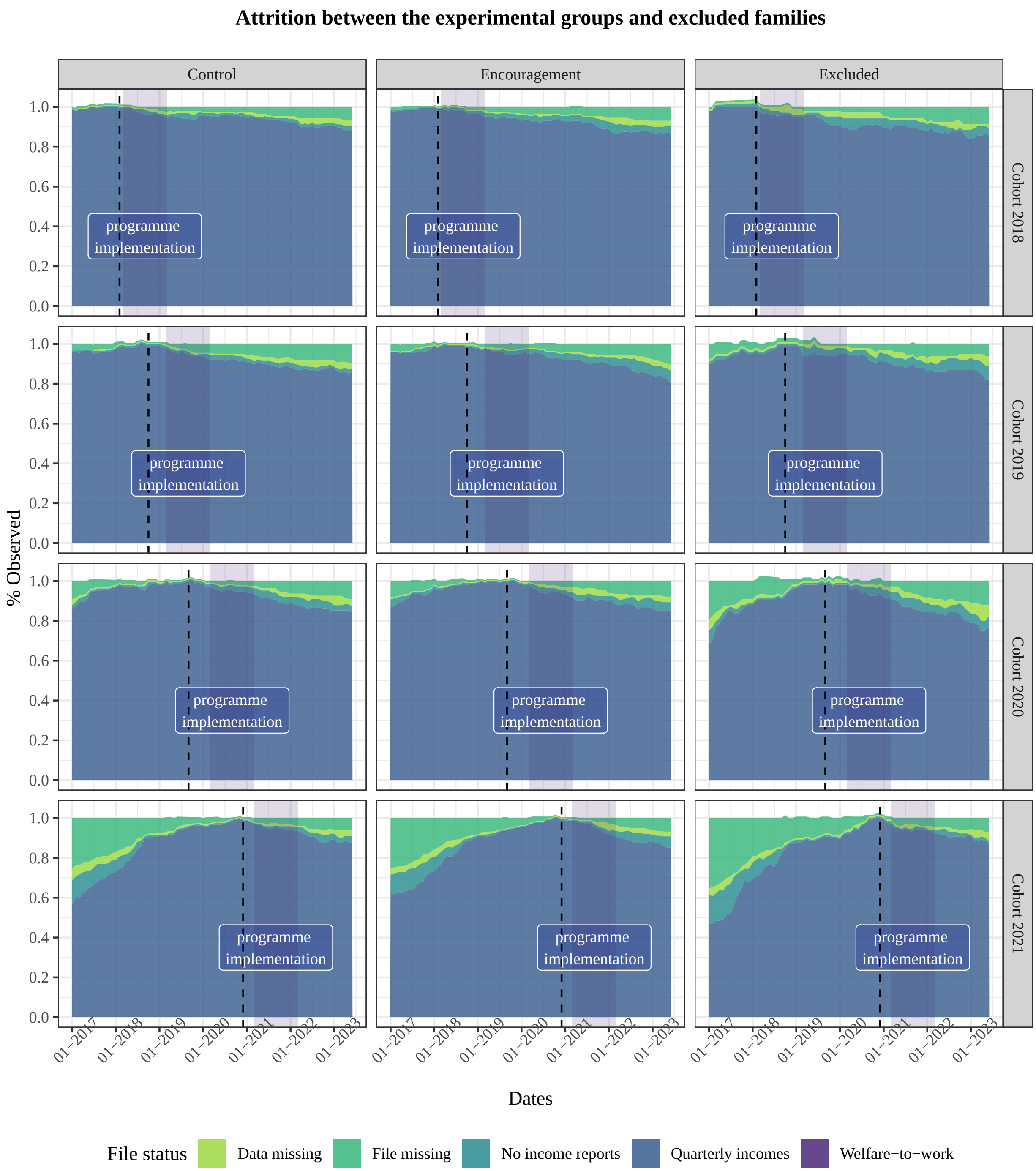
Table B.4: Definition of main outcome variables

Outcomes	Variables	Variable names	Description
Main outcomes	Labour Income >0	income_pos	*Dummy for positive labour income of the sampled parent. Built from quarterly reports for RSA and/or PA (MTACMMEX|MTACMONX)*
	Receive Rsa	out_RSA	*Dummy for receiving RSA (at least € 15 paid)*
	Monthly Total Household's Incomes	incomeRUC	*Total disposable incomes of the household, including social transfers and spouse incomes. Computed by the national family allowance fund and actualised using the INSEE CPI for bottom quintile of the income distribution.*
	Monthly Household's Incomes Per Cu	RUC	*Total household incomes weighted by family size and structure. First adult has weight 1, any other adult or child over 14 has weight .5, additional children under 14 have .3 weights. Single parents have an additional .2 weight. Computed by the national family allowance fund and actualised using the INSEE CPI for bottom quintile of the income distribution.*
	Monthly Total Social Transfers	MTPRESVE	*Total transfers from the Family allowance fund. Computed by the national family allowance fund actualised using the INSEE CPI for bottom quintile of the income distribution.*
Alternative measurements	'In Work'	Employment	*Dummy that equals one if parents 'main' occupation is 'in work' according to INSEE 7 categories employment classification. Coded from ACTRESPD, imputed by social workers when files are updated. Low reliability on the timing*
	Incomes Above Rsa Threshold	out_RSAtight	*Dummy for exiting RSA because earnings are above the eligibility threshold. Computed from motives for RSA ineligibility or payment cancelling.*
	Individual Earnings	INCOME	*Individual earnings of the sampled parent computed from quarterly reports*

B.III Attrition

A) Descriptive statistics

Figure B.12: Share of income data available across time and cohorts

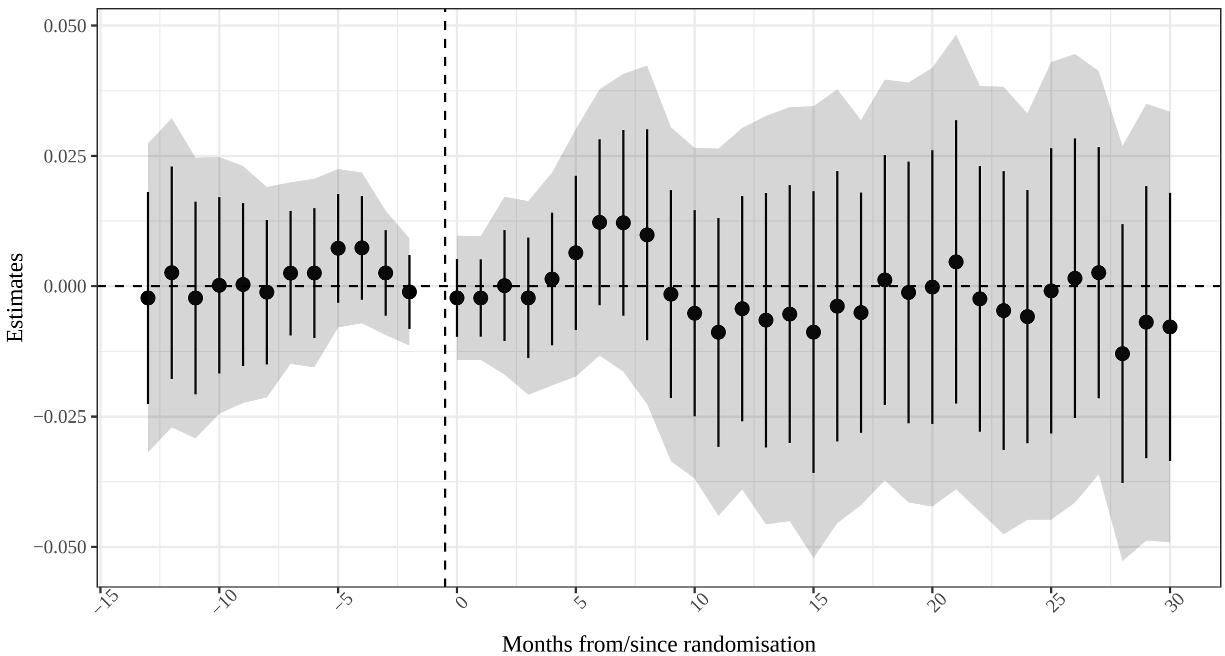


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Proportions of the baseline population observed or not at each date
in the experimental group and excluded families.

B) Estimating the intention-to-treat effect on the probability of reporting quarterly incomes

Figure B.13: Event study of the effect of encouragement on availability of incomes' data

Effect of encouragement on the probability of reporting quarterly incomes



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable equals 1 if we observe quartely incomes the corresponding month, 0 otherwise.
Event study with block x cohort x relative month fixed effect and centered encouragement by block−propensity scores.
Error bars, indicates 95 % CI using cluster−robust standard errors at the cohort level.
Shades indicates 95%CI adjusting for the FWER.

B.IV Other variables used in the analysis

Table B.5: Definition and label for variables used as covariates

Variables	Variable names	Description
Age	AGE	*Parent's age at the month of randomisation. Computed from Birth date in the ALLSTAT data.*
Age youngest child	youngest	*Age of the youngest child computed from ALLSTAT data.*
Age oldest child	oldest	*Age of the oldest child computed from ALLSTAT data.*
Gender	Sexe	*Gender as defined by social security numbers, original renamed from ALLSTAT data.*
N Children under 2	NBEN0A2C	*Number of children under 2 year-old in december of the year, computed by CAF.*
N Children 3 to 5	NBEN3A5C	*Number of children 3 to 5 years-old in december of the year, computed by CAF.*
STI assessment	SESIDummy	*Dummy for social workers assessment of the family's appropriatness for the program. 1 for favourable assessment, 0 for reserved.*
French citizenship	French	*Dummy for French cityzenship from ALLSTAT*
Family allowance	AF	*Dummy for receiving 'Allocations familiales', constructed from AFVERS*
Family supplement	CF	*Dummy for receiving 'Complément familiale', constructed from CFVERS*
Housing benefit	APL	*Dummy for receiving one of the three housing benefit, constructed from DROAIDLO*
Family support allowance	ASF	*Dummy for receiving 'Allocation de soutien familiale', constructed from ASFVERS*
Child support	TPA	*Dummy for receiving 'pension alimentaire',original renamed from ALLSTAT*
Education	School3	*Level of education grouped in 3 values (High/Low/Unknown): High := for High school degree , Higher education and Vocational degree ; Low:= Middle school diploma , No education , Other training or Some Middle school.*
N consumption units	NBUC	*Number of consumption units computed by Cnaf. 1 unit of for the first adult, 0.5 per additional adult or child aged 14 and over, 0.3 per child under 14 and 0.2 for a single-parent family.*

Sources: ALLSTAT. Data preparation files are available upon request.

Table B.5: Definition and label for variables used as covariates

Variables	Variable names	Description
Early childhood allowance	PAJE	*Dummy for receiving 'Préstation d'accueil du jeune enfant', constructed from PAJEVERS*

Sources: ALLSTAT. Data preparation files are available upon request.

B.V Balance check at the time of random assignment

Table B.6: Balance of main variables of interest the month before randomisation

	Control (N=828)		Encouragement (N=843)		Diff. in Means	Std. Error
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*		
Share labour income >0	0.08	0.28	0.08	0.27	-0.01	0.02
Share receive RSA	0.98	0.15	0.98	0.14	0.00	0.01
Mean monthly total household's incomes	1394.69	493.21	1398.98	497.86	4.37	24.08
Mean monthly household's incomes per CU	708.59	150.11	712.19	157.06	3.70	11.51
Mean monthly total social transfers	1297.67	497.91	1292.00	481.42	-7.10	17.99
Mean yearly taxable income N-2	1424.11	2871.59	1599.39	3315.34	178.27	206.54
Favourable assessment	0.66	0.47	0.69	0.46	0.03	0.03
Mean distance (km) to the programme	3.32	1.86	3.49	2.01	0.17	0.12
Share French	0.81	0.39	0.84	0.37	0.03	0.02
Share higher education	0.51	0.50	0.53	0.50	0.02	0.03
Share lower education	0.25	0.43	0.24	0.43	0.00	0.03
Share unknown education	0.24	0.43	0.23	0.42	-0.01	0.02
Mean age	36.04	7.95	36.14	7.75	0.08	0.46
Mean age youngest child	7.13	5.56	7.16	5.49	0.03	0.34
Mean age oldest child	11.41	6.28	11.32	6.12	-0.11	0.34
Share with children under 2	0.31	0.46	0.29	0.45	-0.02	0.03
Share with children 3 to 5	0.33	0.47	0.32	0.47	-0.01	0.03
Share with one child over 16	0.32	0.47	0.30	0.46	-0.03	0.03
Share receive family allowance	0.57	0.50	0.56	0.50	-0.01	0.01
Share receive family supplement	0.17	0.38	0.17	0.38	0.00	0.03
Share receive housing benefit	0.89	0.31	0.88	0.32	-0.01	0.02

* = p<.1, ** = p<.05, *** = p<.01
Sources: ALLSTAT, cohorts 2018 to 2021 one month before randomisation.
Notes : mean and mean differences are weighted within-block averages.
Standard errors account for block randomisation.

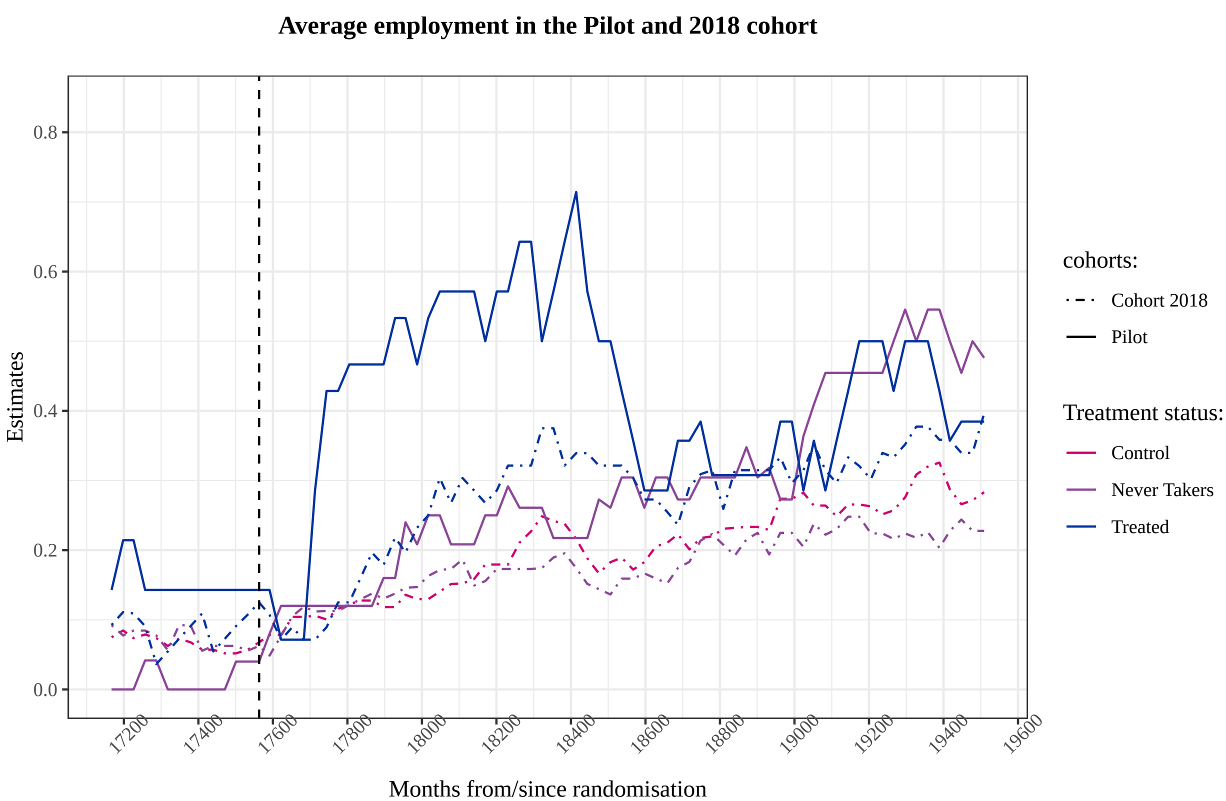
Table B.6: Balance of main variables of interest the month before randomisation

	Control (N=828)		Encouragement (N=843)		Diff. in Means	Std. Error
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*		
Share receive family support allowance	0.65	0.48	0.64	0.48	-0.01	0.02
Share Receive child support	0.21	0.40	0.21	0.40	0.00	0.02
Share receive Early childhood allowance	0.31	0.46	0.29	0.45	-0.02	0.03

* = p<.1, ** = p<.05, *** = p<.01
Sources: ALLSTAT, cohorts 2018 to 2021 one month before randomisation.
Notes : mean and mean differences are weighted within-block averages.
Standard errors account for block randomisation.

B.VI Comparison of the outcomes of the 2018 cohort and the pilot group

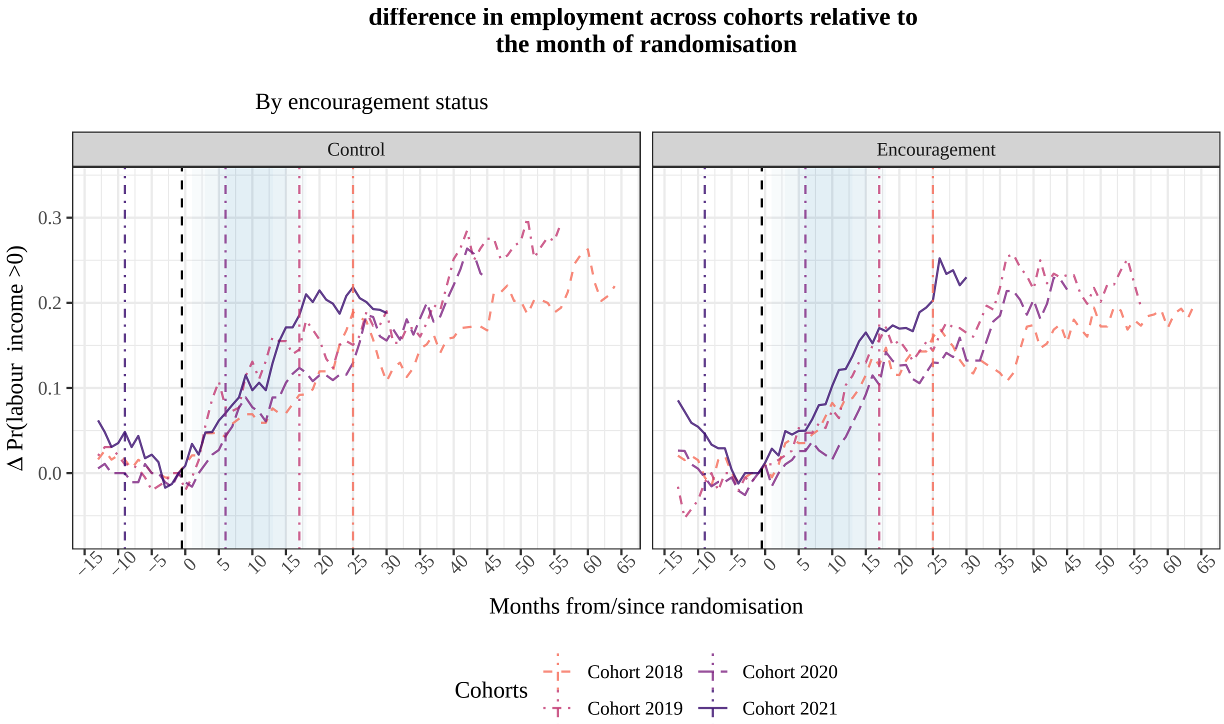
Figure B.14: Evolution of the share of parents with positive labour income in the pilot and first cohort



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable equals 1 when the person has positive labour income and 0 otherwise.

B.VII Comparison of average changes in employment in relative time across cohorts and treatment arms

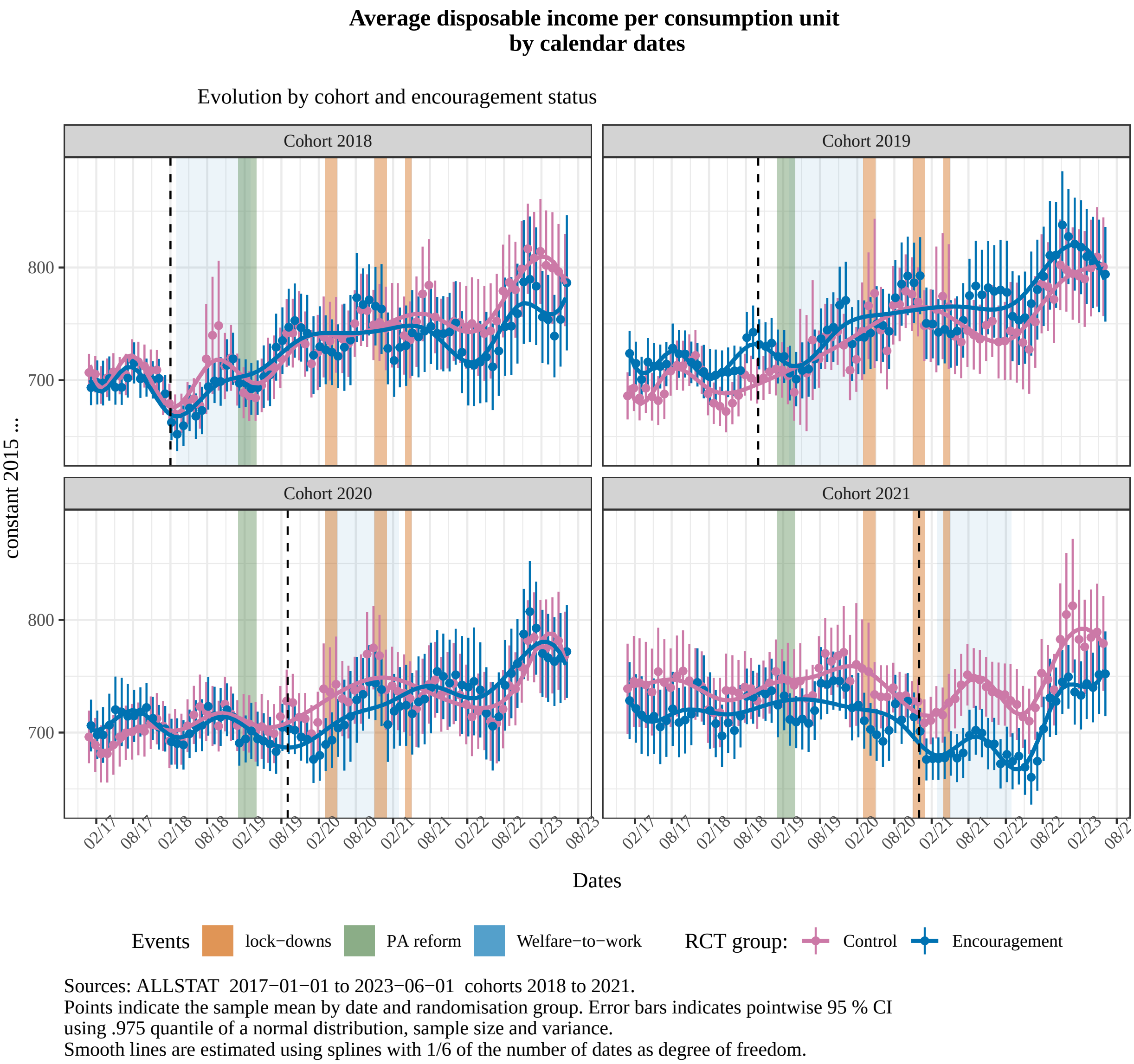
Figure B.15: Never more than 25 pp increase in employment in all 4 cohorts



Sources: ALLSTAT 2017−01−01 to 2023−06−01
The dependent variable is the long difference between employment at month m and employment at the time of randomisation.
Lines indicates sample means color codded by cohort.
The dash−dot lines indicate the month of the first lock−down for each cohort.

B.VIII Comparison of average disposable income per consumption units across cohorts and treatment arms

Figure B.16: Average disposable income per consumption unit over the period



C Additional estimates of the effect of the programme

C.I Aggregated treatment effect on employment

Table C.7: Aggregated estimates on labour market maret participation

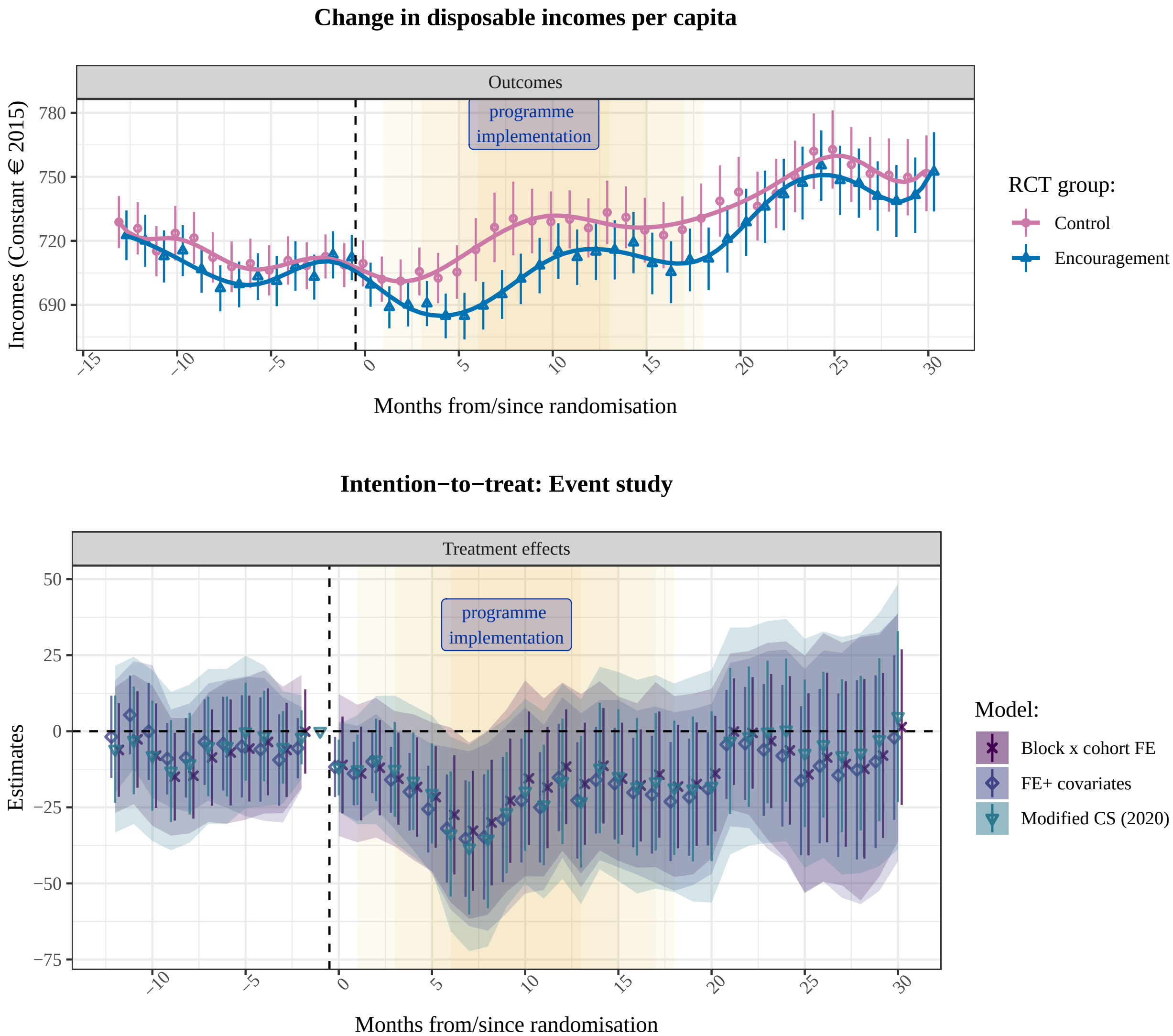
	Mean control	**OLS**		**TSLS**	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	0.085*** (0.012)	-0.009 (0.013)	-0.006 (0.009)	-0.024 (0.034)	-0.017 (0.023)
	[0.06, 0.11]	[-0.04, 0.02]	[-0.03, 0.02]	[-0.11, 0.06]	[-0.08, 0.04]
	adj.p.val. = 0.000	*adj.p.val. = 0.927*	*adj.p.val. = 0.945*	*adj.p.val. = 0.927*	*adj.p.val. = 0.944*
[0 ; 6 [	0.113*** (0.012)	-0.014 (0.012)	-0.015 (0.010)	-0.036 (0.030)	-0.039 (0.026)
	[0.08, 0.14]	[-0.04, 0.02]	[-0.04, 0.01]	[-0.11, 0.04]	[-0.11, 0.03]
	adj.p.val. = 0.000	*adj.p.val. = 0.626*	*adj.p.val. = 0.477*	*adj.p.val. = 0.627*	*adj.p.val. = 0.473*
[6 ; 12 [	0.164*** (0.016)	-0.029** (0.014)	-0.036** (0.014)	-0.073** (0.034)	-0.093*** (0.035)
	[0.13, 0.20]	[-0.06, 0.01]	[-0.07, 0.00]	[-0.16, 0.01]	[-0.18, 0.00]
	adj.p.val. = 0.000	*adj.p.val. = 0.140*	*adj.p.val. = 0.051*	*adj.p.val. = 0.128*	*adj.p.val. = 0.039*
[12 ; 18 [	0.203*** (0.020)	-0.006 (0.016)	-0.013 (0.016)	-0.016 (0.042)	-0.032 (0.042)
	[0.16, 0.25]	[-0.05, 0.04]	[-0.05, 0.03]	[-0.12, 0.09]	[-0.14, 0.07]
	adj.p.val. = 0.000	*adj.p.val. = 0.997*	*adj.p.val. = 0.927*	*adj.p.val. = 0.997*	*adj.p.val. = 0.927*
[18 ; 24 [	0.235*** (0.021)	-0.009 (0.019)	-0.020 (0.017)	-0.022 (0.047)	-0.052 (0.042)
	[0.18, 0.29]	[-0.06, 0.04]	[-0.06, 0.02]	[-0.14, 0.10]	[-0.16, 0.06]
	adj.p.val. = 0.000	*adj.p.val. = 0.989*	*adj.p.val. = 0.675*	*adj.p.val. = 0.989*	*adj.p.val. = 0.665*
[24 ; 30 [	0.260*** (0.020)	-0.008 (0.020)	-0.014 (0.019)	-0.021 (0.049)	-0.036 (0.048)
	[0.21, 0.31]	[-0.06, 0.04]	[-0.06, 0.03]	[-0.15, 0.10]	[-0.16, 0.09]
	adj.p.val. = 0.000	adj.p.val. = 0.994	adj.p.val. = 0.940	adj.p.val. = 0.994	adj.p.val. = 0.940
Num.Obs.	28125	56749	56749	56749	56749
R2	0.027	0.108	0.280	0.106	0.274
R2 Adj.	0.027	0.067	0.245	0.064	0.239
Covariates			X		X

* $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$ using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods. Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

C.II Dynamic intention to treat on income per consumption unit

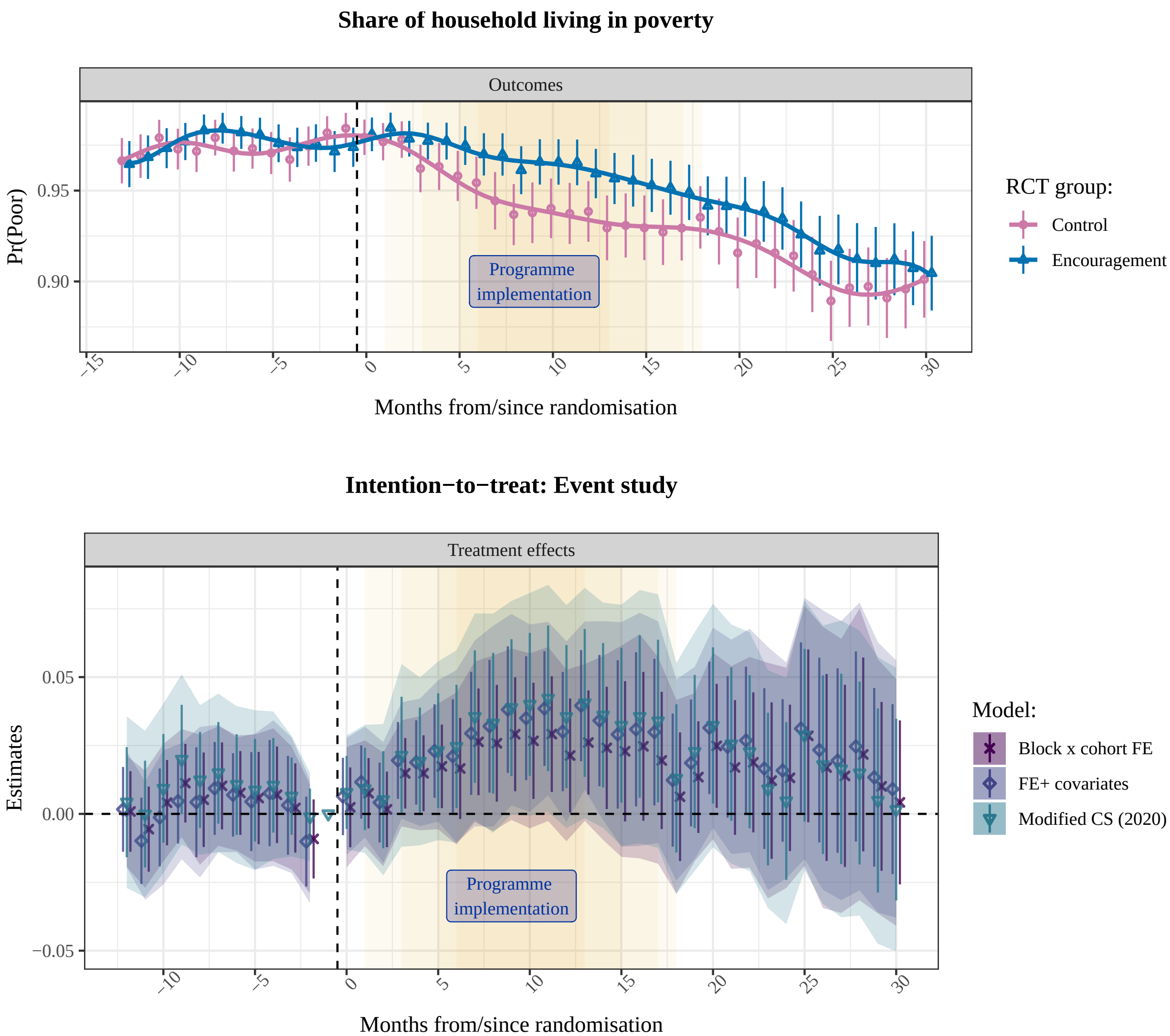
Figure C.17: Dynamic Effects of the programme on incomes per capita



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the disposable income per consumption unit.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS or modified Callaway Sant' Anna (2020).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

C.III Dynamic intention to treat on the risk of poverty

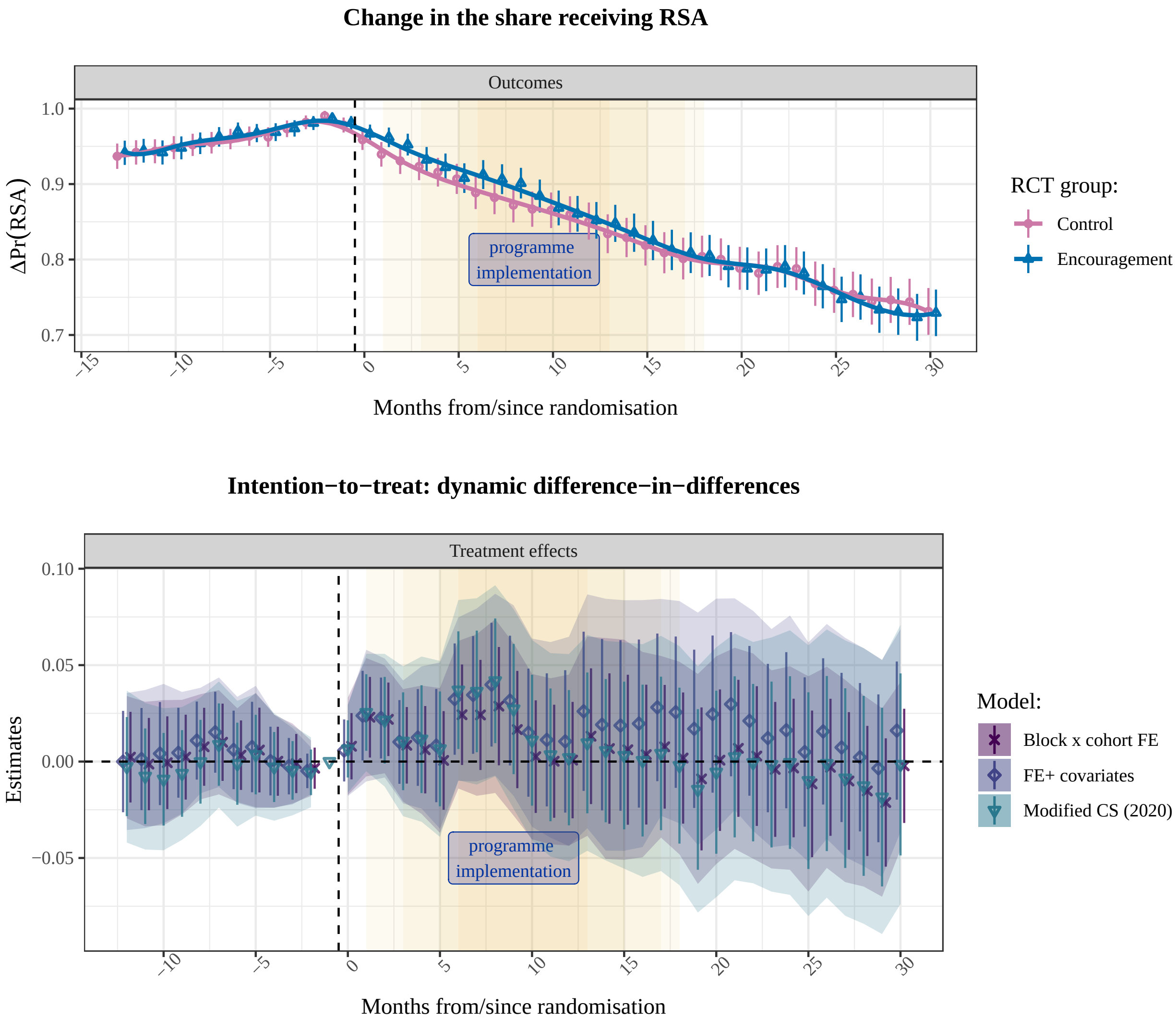
Figure C.18: Dynamic Effects of the programme on the probability of living in poverty



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for disposable income per capita lower than the poverty line.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS or modified Callaway Sant' Anna (2020).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level .
− Shades indicates 95%CI adjusting for the FWER with Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

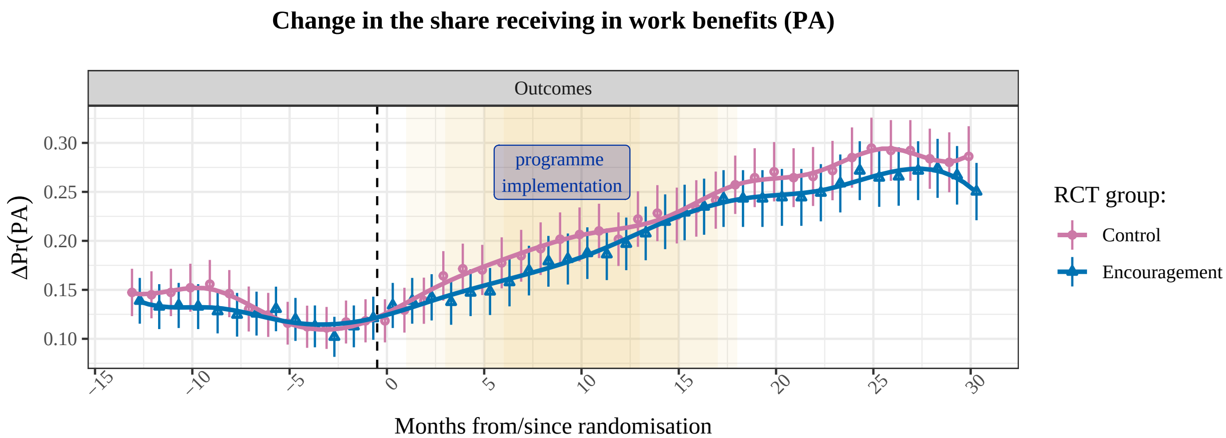
C.IV Effect of the programme on RSA and PPA take-up

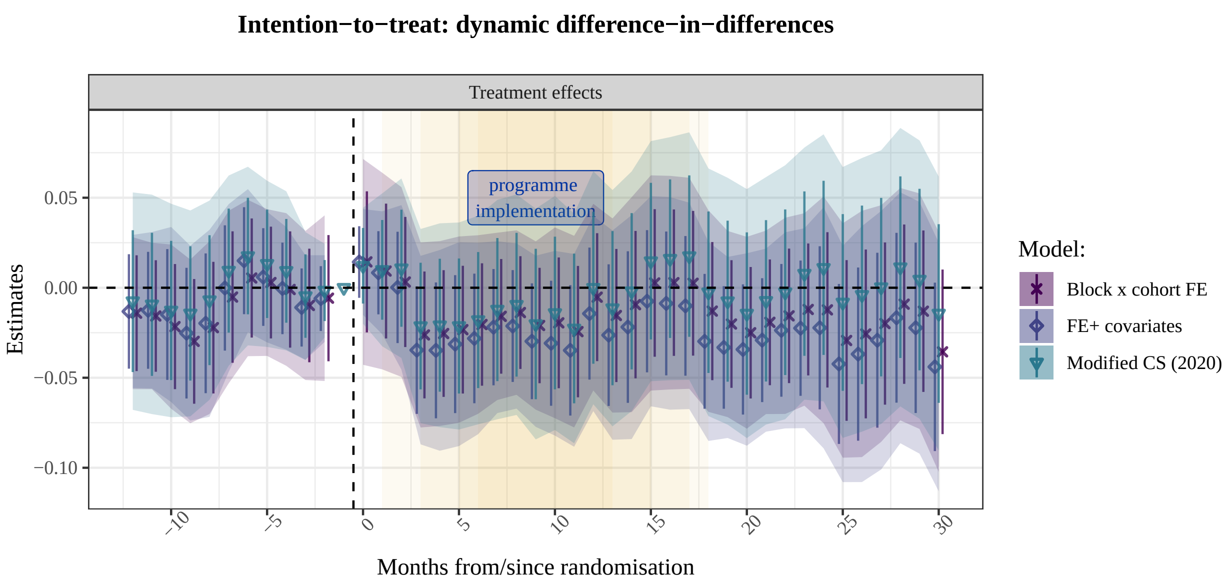
Figure C.19: Intention-to-treat effects of the programme on the probability of receiving RSA payments



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for receiving RSA.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

Figure C.20: Intention-to-treat effects of the programme on the probability of receiving in work benefits (PA)



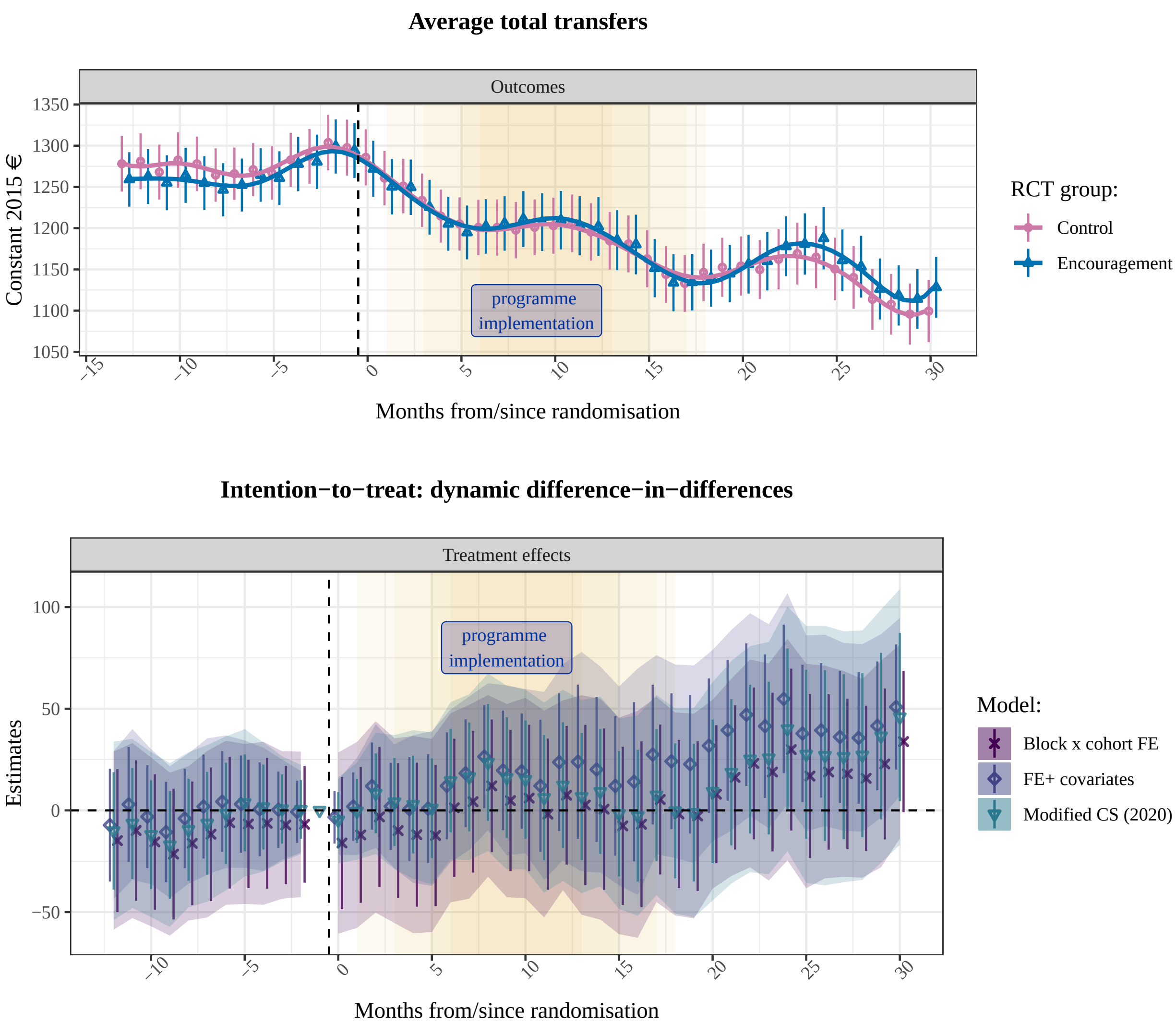


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for receiving PA, the in−work benefit.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

C.V Effects of the programme on total transfers from the Family allowance fund

A) Dynamic intention to treat on total amount of cash transfers

Figure C.21: Intention-to-treat effects of the programme on total amount of transfers from the Family allowance fund



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the total amount of cash transfers, in constant 2015 symbol("\xa0") .
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

B) Aggregated effect on total social transfers

Table C.8: Aggregated effects of the programme on total transfers from the family allowance fund

	Mean control	OLS		TSLS	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	1280.2*** (52.1)	-7.5 (14.5)	1.5 (9.5)	-19.3 (37.7)	4.0 (24.7)
	[1168.5, 1391.9]	[-43.1, 28.2]	[-22.4, 25.5]	[-112.0, 73.4]	[-58.2, 66.3]
	adj.p.val. = 0.000	*adj.p.val. = 0.964*	*adj.p.val. = 1.000*	*adj.p.val. = 0.965*	*adj.p.val. = 1.000*
[0 ; 6 [	1242.3*** (51.3)	-10.9 (16.1)	2.4 (9.0)	-28.0 (41.8)	6.3 (23.3)
	[1132.6, 1352.1]	[-50.5, 28.6]	[-20.3, 25.2]	[-130.8, 74.7]	[-52.3, 64.9]
	adj.p.val. = 0.000	*adj.p.val. = 0.898*	*adj.p.val. = 0.999*	*adj.p.val. = 0.902*	*adj.p.val. = 0.999*
[6 ; 12 [	1202.4*** (50.0)	4.5 (16.6)	18.0 (12.8)	11.4 (42.4)	46.2 (32.7)
	[1095.3, 1309.5]	[-36.2, 45.2]	[-14.1, 50.0]	[-93.0, 115.8]	[-36.0, 128.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.998*	*adj.p.val. = 0.476*	*adj.p.val. = 0.998*	*adj.p.val. = 0.474*
[12 ; 18 [	1167.4*** (50.4)	0.3 (18.3)	20.2 (16.8)	0.7 (47.0)	51.9 (43.0)
	[1059.4, 1275.4]	[-44.7, 45.3]	[-22.1, 62.5]	[-114.8, 116.3]	[-56.2, 160.0]
	adj.p.val. = 0.000	*adj.p.val. = 1.000*	*adj.p.val. = 0.623*	*adj.p.val. = 1.000*	*adj.p.val. = 0.620*
[18 ; 24 [	1155.9*** (49.5)	10.2 (17.5)	34.5** (16.1)	26.2 (44.5)	88.2** (41.0)
	[1049.8, 1261.9]	[-32.6, 53.1]	[-5.9, 74.8]	[-83.3, 135.6]	[-15.1, 191.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.939*	*adj.p.val. = 0.123*	*adj.p.val. = 0.939*	*adj.p.val. = 0.123*
[24 ; 30 [	1128.9*** (48.9)	20.4 (17.8)	41.0*** (15.2)	51.9 (45.5)	104.4** (39.6)
	[1024.1, 1233.7]	[-23.3, 64.0]	[2.8, 79.2]	[-60.1, 163.9]	[4.7, 204.1]
	adj.p.val. = 0.000	adj.p.val. = 0.586	adj.p.val. = 0.031	adj.p.val. = 0.593	adj.p.val. = 0.036
Num.Obs.	28963	58486	58486	58486	58486
R2	0.011	0.525	0.721	0.524	0.718
R2 Adj.	0.010	0.503	0.708	0.502	0.705
Covariates			X		X

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods.

Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

C) Aggregated effect on the probability of living in poverty

Table C.9: Aggregated effects of the programme on poverty

	Mean control	**OLS**		**TSLS**	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	0.97*** (0.01)	0.00 (0.01)	0.00 (0.01)	0.01 (0.02)	0.01 (0.02)
	[0.96, 0.99]	[-0.01, 0.02]	[-0.01, 0.02]	[-0.03, 0.05]	[-0.04, 0.06]
	adj.p.val. = 0.000	*adj.p.val. = 0.980*	*adj.p.val. = 0.988*	*adj.p.val. = 0.981*	*adj.p.val. = 0.992*
[0 ; 6 [	0.97*** (0.00)	0.01* (0.01)	0.01* (0.01)	0.03* (0.01)	0.03** (0.02)
	[0.96, 0.98]	[0.00, 0.02]	[0.00, 0.03]	[-0.01, 0.06]	[-0.01, 0.07]
	adj.p.val. = 0.000	*adj.p.val. = 0.296*	*adj.p.val. = 0.211*	*adj.p.val. = 0.276*	*adj.p.val. = 0.206*
[6 ; 12 [	0.94*** (0.01)	0.03*** (0.01)	0.03*** (0.01)	0.07*** (0.02)	0.08*** (0.02)
	[0.92, 0.96]	[0.00, 0.05]	[0.01, 0.05]	[0.01, 0.12]	[0.02, 0.14]
	adj.p.val. = 0.000	*adj.p.val. = 0.016*	*adj.p.val. = 0.007*	*adj.p.val. = 0.013*	*adj.p.val. = 0.006*
[12 ; 18 [	0.93*** (0.01)	0.02** (0.01)	0.03*** (0.01)	0.06** (0.02)	0.08*** (0.03)
	[0.91, 0.95]	[0.00, 0.05]	[0.00, 0.06]	[0.00, 0.12]	[0.01, 0.15]
	adj.p.val. = 0.000	*adj.p.val. = 0.091*	*adj.p.val. = 0.016*	*adj.p.val. = 0.078*	*adj.p.val. = 0.013*
[18 ; 24 [	0.92*** (0.01)	0.02 (0.01)	0.02* (0.01)	0.04 (0.03)	0.05* (0.03)
	[0.90, 0.95]	[-0.01, 0.04]	[-0.01, 0.05]	[-0.03, 0.10]	[-0.02, 0.12]
	adj.p.val. = 0.000	*adj.p.val. = 0.477*	*adj.p.val. = 0.254*	*adj.p.val. = 0.459*	*adj.p.val. = 0.245*
[24 ; 30 [	0.90*** (0.01)	0.02 (0.01)	0.02 (0.01)	0.04 (0.03)	0.05 (0.03)
	[0.86, 0.93]	[-0.02, 0.05]	[-0.01, 0.05]	[-0.04, 0.13]	[-0.03, 0.13]
	adj.p.val. = 0.000	adj.p.val. = 0.666	adj.p.val. = 0.469	adj.p.val. = 0.657	adj.p.val. = 0.475
Num.Obs.	28700	57927	57927	57927	57927
R2	0.013	0.067	0.095	0.065	0.088
R2 Adj.	0.012	0.025	0.051	0.022	0.044
Covariates			X		X

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

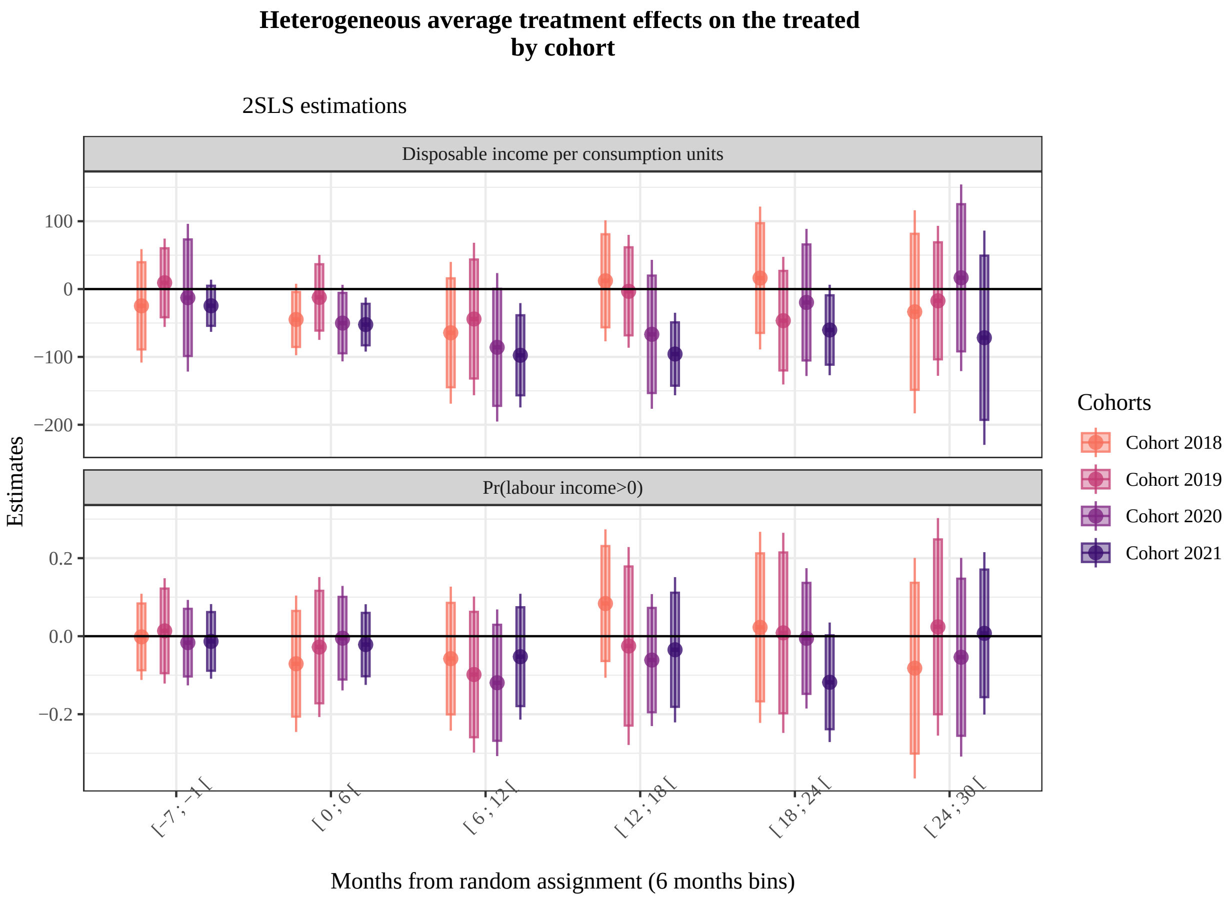
Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods.

Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

D Estimations of heterogeneous treatment effects

D.I Treatment effect heterogeneity on employment by cohort

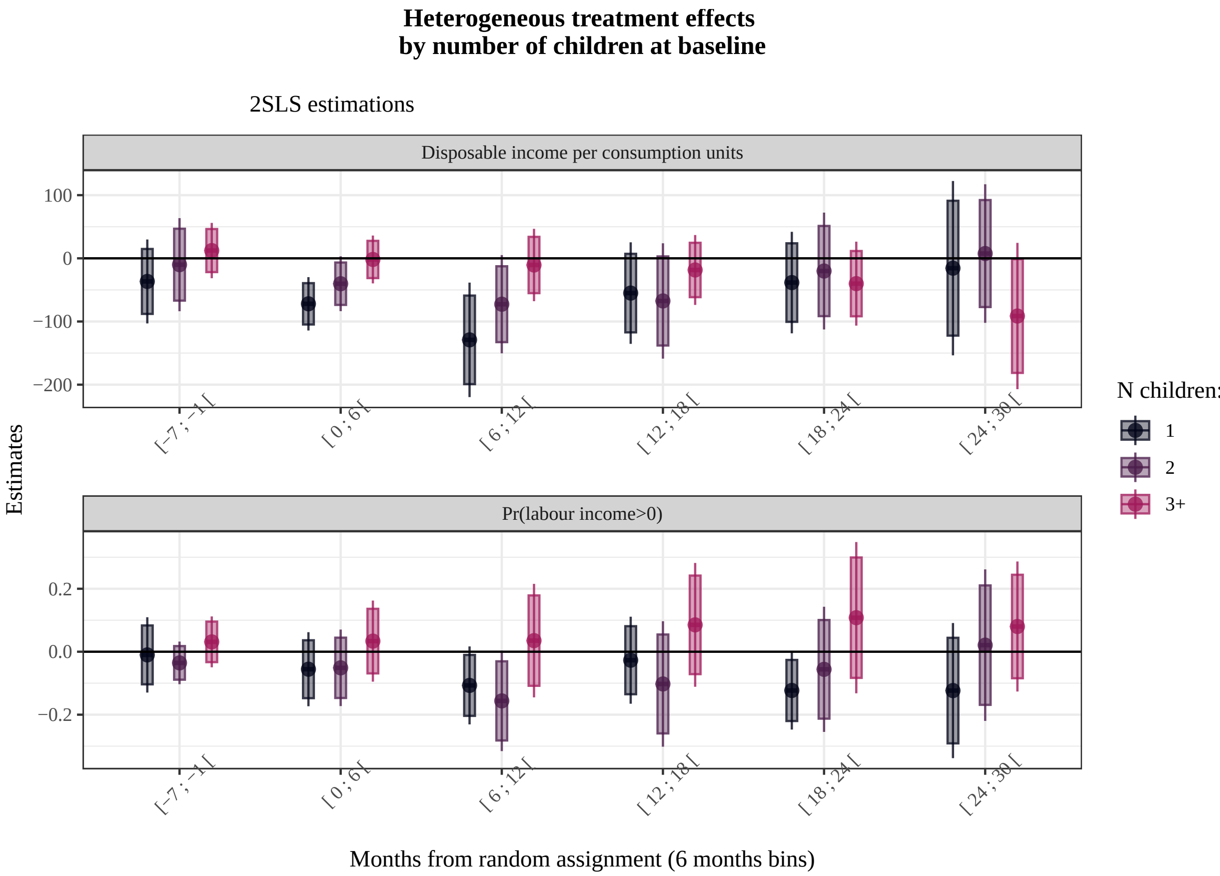
Figure D.22: Heterogeneity on effect on labour market participation across cohorts



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with cohort x participation x period dummies
instrumented by the demeaned instrument x cohort x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.II Treatment effect heterogeneity by number of children at baseline

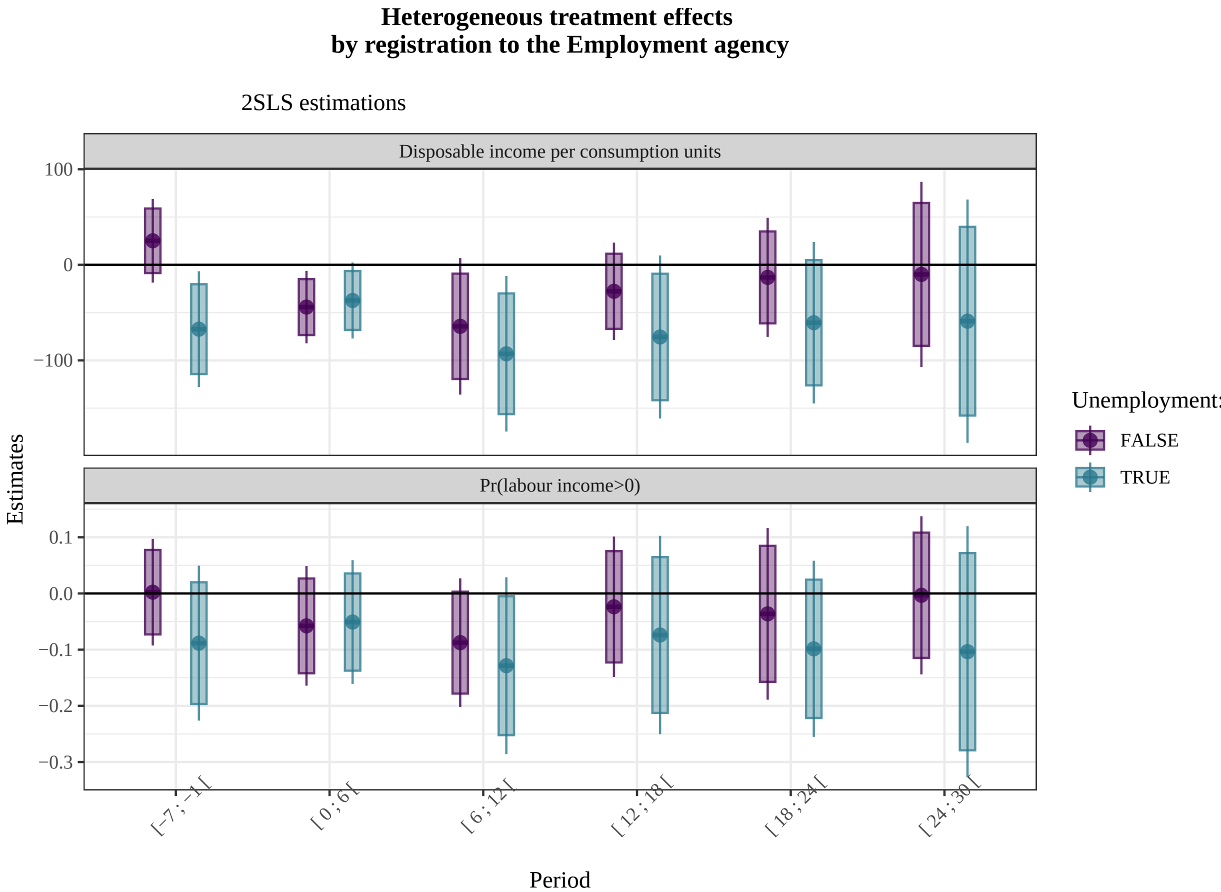
Figure D.23: Heterogenous treatment effects on disposable incomes per capita by number of children



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with children x participation x period dummies instrumented by the demeaned instrument x children x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.III Treatment effect heterogeneity on disposable incomes by registration to the Employment agency

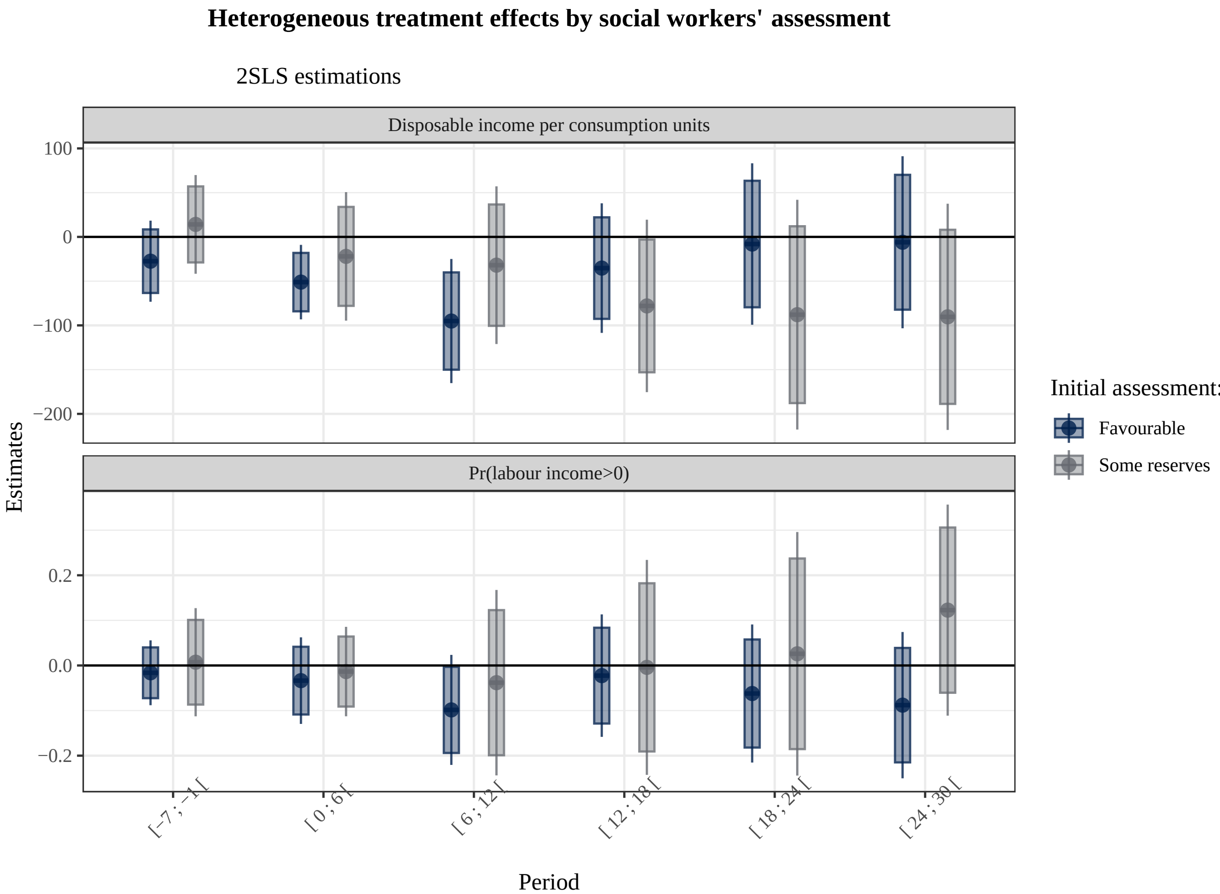
Figure D.24: Heterogenous treatment effects on employment by unemployment status



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with unemployment x participation x period dummies
instrumented by the demeaned instrument x unemployment x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.IV Treatment effect heterogeneity by social workers’ initial assessment

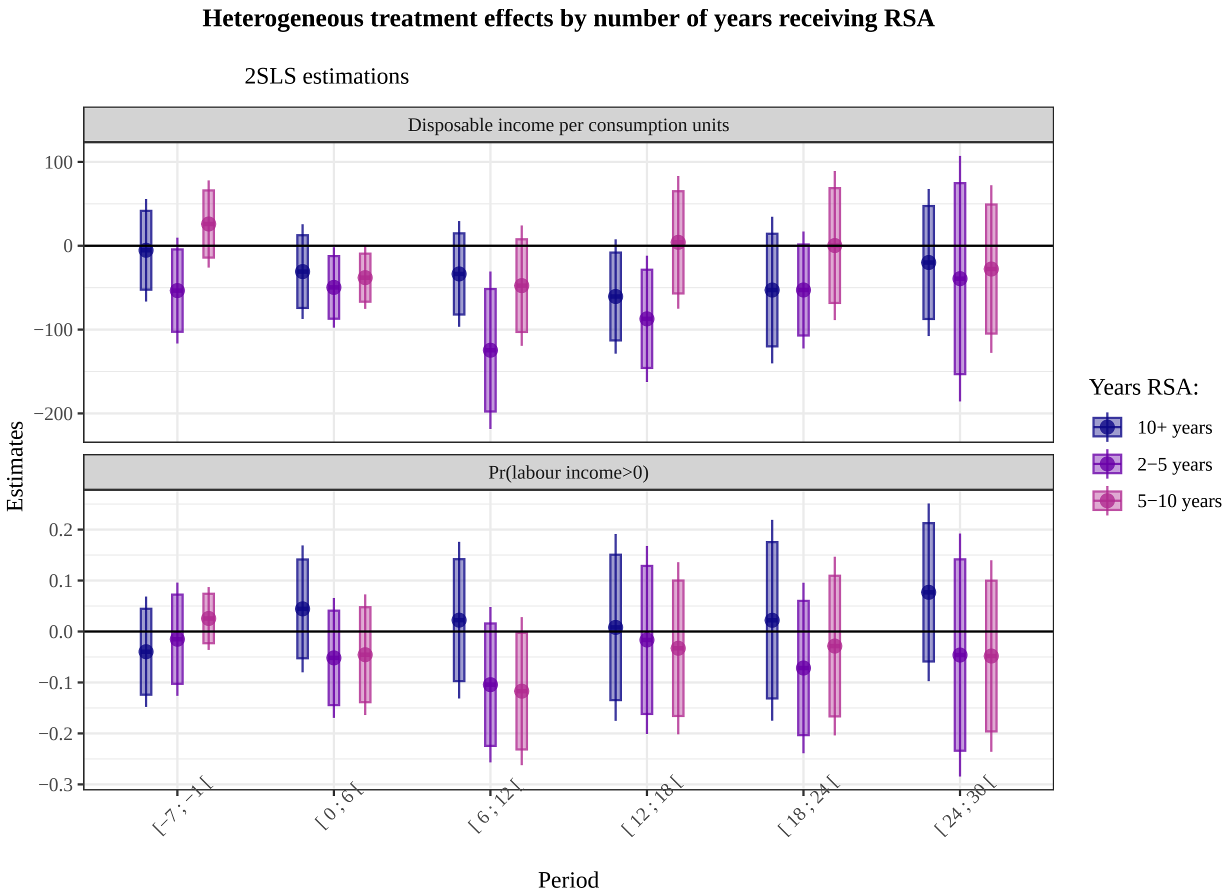
Figure D.25: Heterogenous treatment effects by social workers’ initial assessment



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with assessment x participation x period dummies instrumented by the demeaned instrument x assessment x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.V Treatment effect heterogeneity by number of years receiving RSA

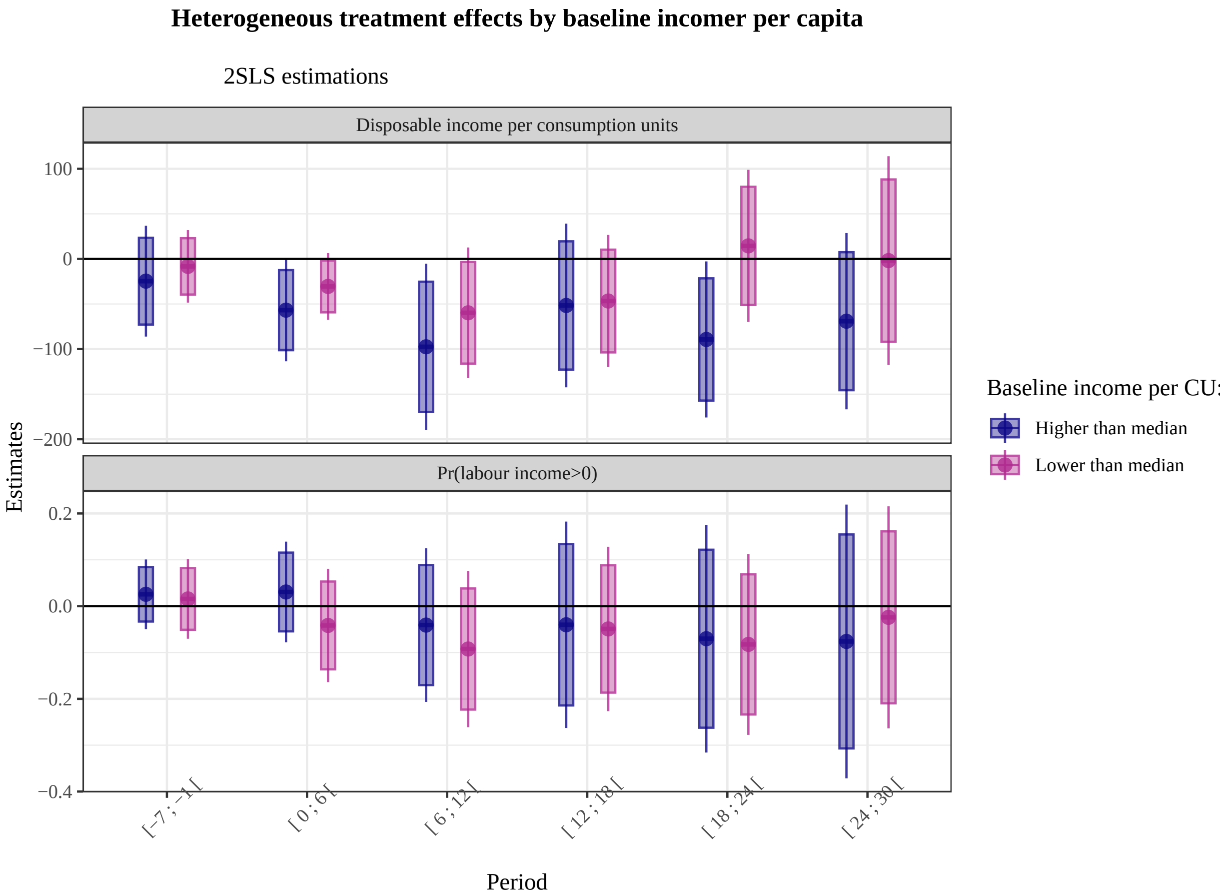
Figure D.26: Heterogenous treatment effects by social workers' initial assessment



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with seniority x participation x period dummies
instrumented by the demeaned instrument x seniority x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.VI Treatment effect heterogeneity by baseline income per capita

Figure D.27: Heterogenous treatment effects by social workers' initial assessment

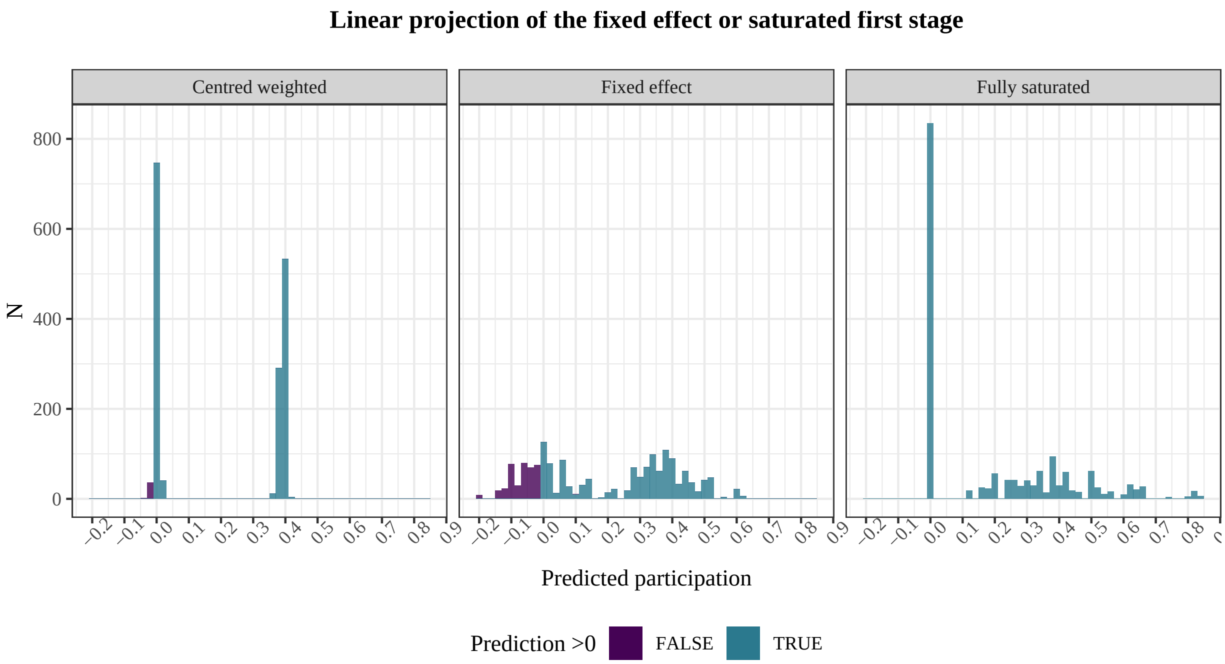


Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with median income x participation x period dummies
instrumented by the demeaned instrument x median income x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

E Robustness checks

E.I Predicted participation from the average first stage

Figure E.28: First stage predictions of the Fixed effect and saturated regressions

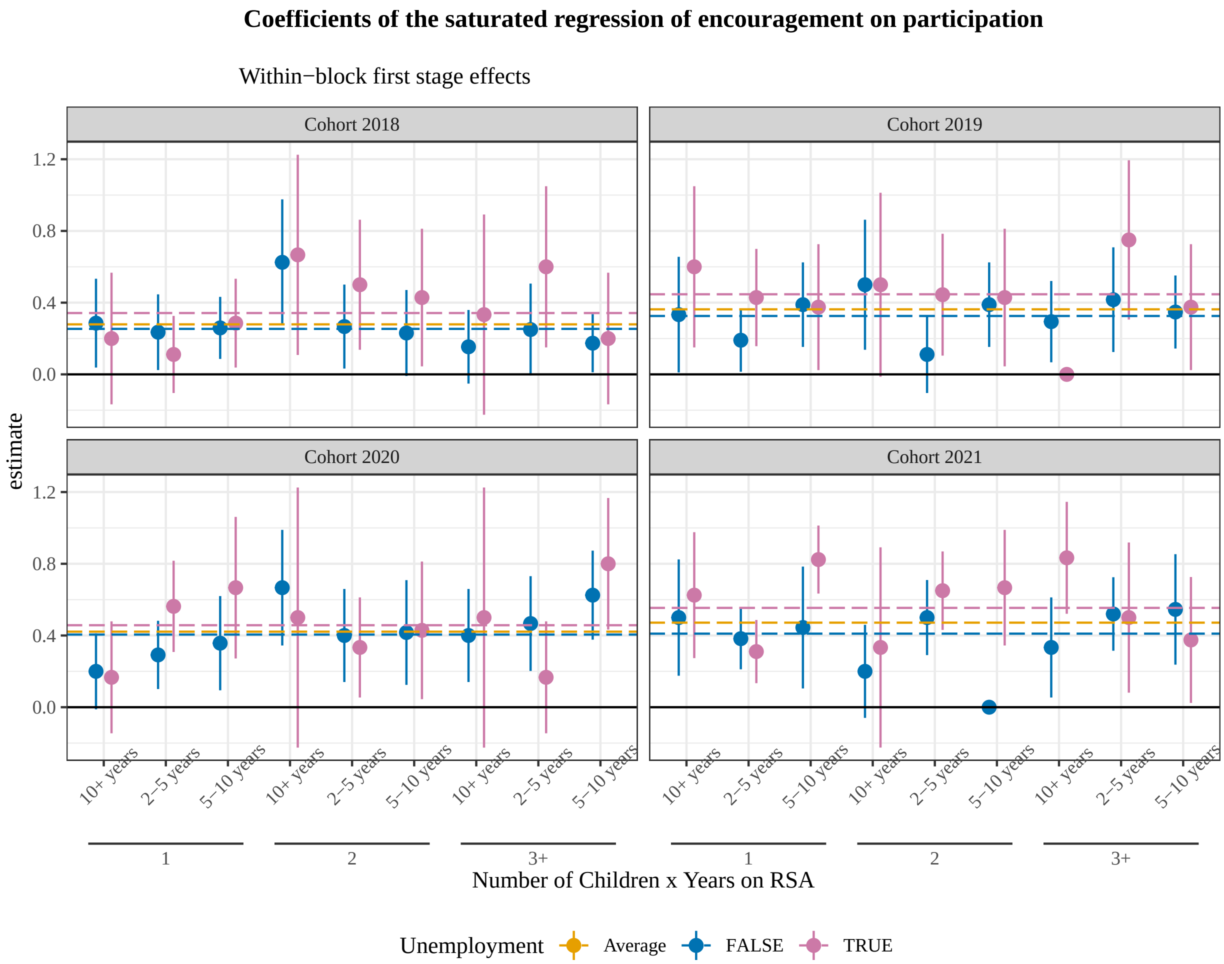


Sources: ALLSTAT, cohorts 2018 to 2021 at baseline.
Notes: Hisograms of the predicted values of the first stage. Binwidth of .02.
All models use inverse instrument propensity score weighting.
The fixed effect model regress participation on encouragement and blocks x cohorts FE.
The saturated model regress participation on the interactions between encouragement
and blocks x cohorts fixed effects.
The Centred weighted model remove block−specific propensity score from the treatment variable.

E.II Coefficients of the fully saturated first stage

In Figure E.29 I plot the coefficients and confidence intervals of the fully saturated first-stage regression; that is, the results of the regression of participation on encouragement × block × cohort fixed effects and block × cohort fixed effects. As this regression is fully saturated with only dummy variables, it estimates the conditional expectation function perfectly, and the coefficient of the interactions estimates the average effect of encouragement on participation in each block.

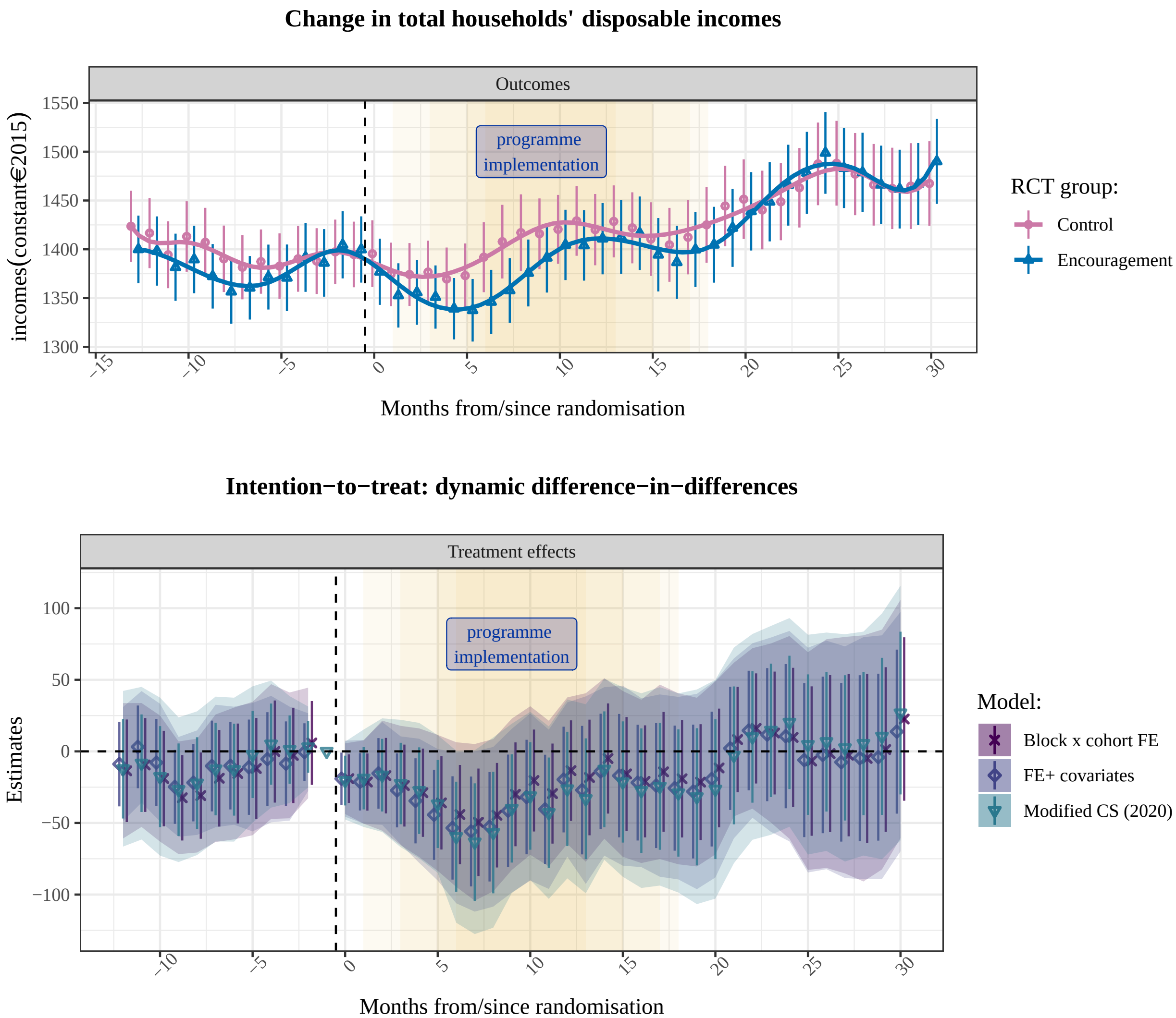
Figure E.29: First stage effect of encouragement on participation in each block



Sources: ALLSTAT cohorts 2018 to 2021 at the month of randomisation.
Notes : Coefficients and 95% confidence intervals of the regression of participation on fixed effects by strata x cohorts and the interaction of encouragement with these fixed effects.
Dashed lines indicate the average effect for each cohort and separate estimates based on unemployment registration.
Error bars are calculated using HC2 heteroskedasticity−robust standard errors.

A) Effects on total household incomes

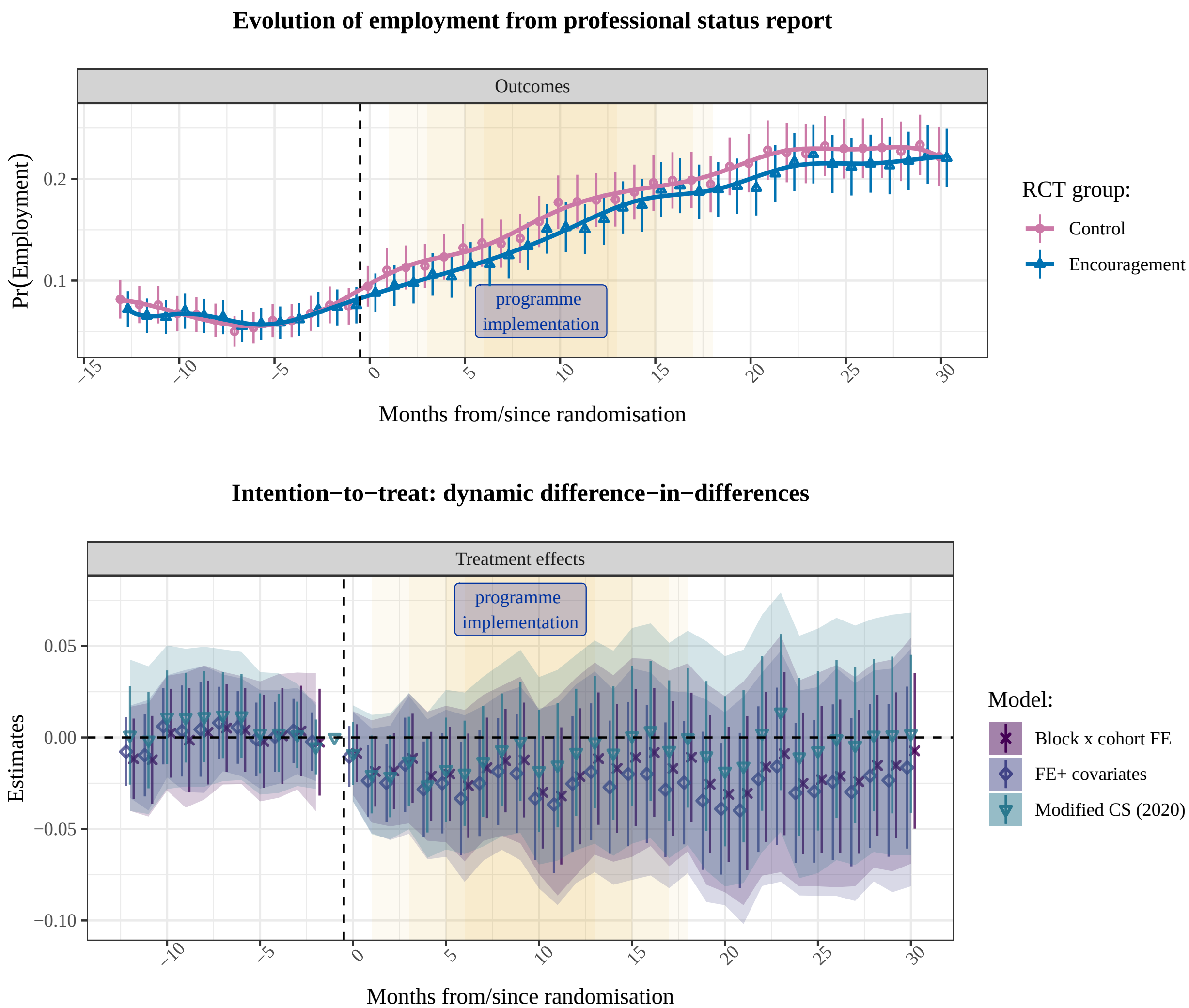
Figure E.30: Intention-to-treat effects of the programme on household disposable incomes estimated by difference-in-differences



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for positive individual labour income.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95% CI adjusting for the FWER using the Holm−Bonferroni correction for OLS models and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

E.III Alternate measure of employment

Figure E.31: Intention-to-treat effects of the programme on household disposable incomes estimated by difference-in-differences



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for being classified as 'Active' in the professional status measure.
This variable is of poor quality as it filled by CAF Agents and ther may be delays.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

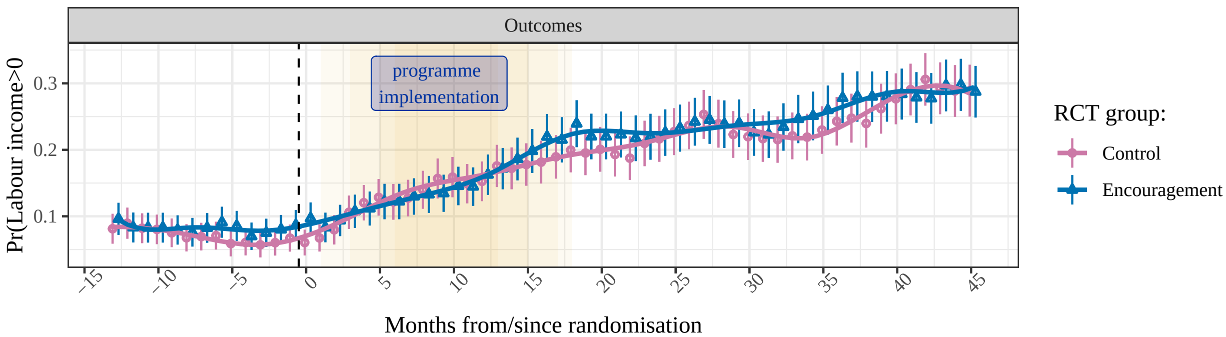
E.IV Estimations for the three first cohorts up to 45 months

A) Intention-to-treat for cohort 2018 to 2020 on labour market participation

Figure E.32: Intention-to-treat effects of the programme on the labour market participation for the three first cohorts up to 45 months.

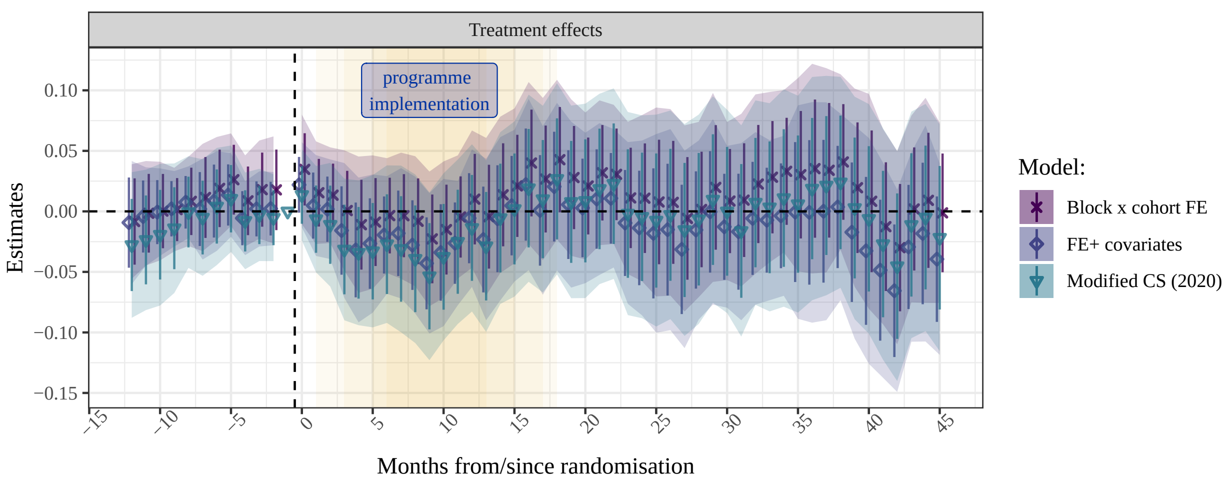
Change in the share with positive labour incomes

Cohort 2018 to 2020 only



Intention−to−treat: dynamic difference−in−differences

Cohort 2018 to 2020 only



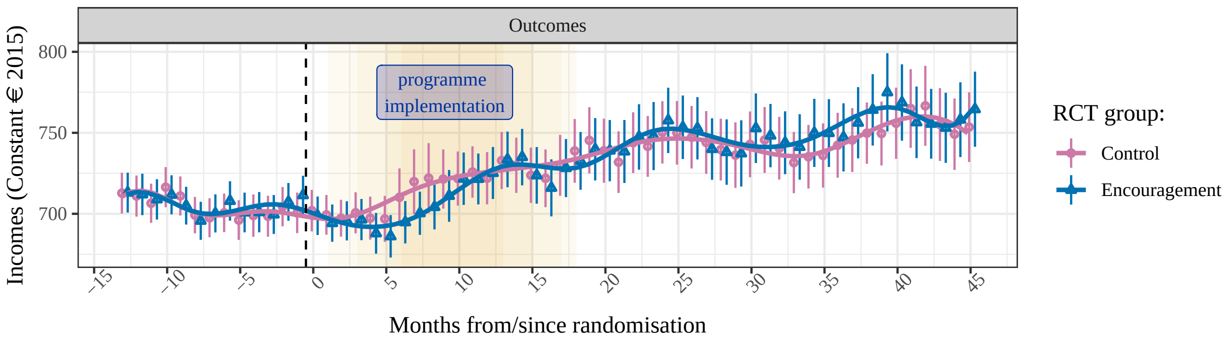
Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for reporting positive labour income.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95% CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

B) Intention-to-treat for cohort 2018 to 2020 on disposable income per capita

Figure E.33: Intention-to-treat effects of the programme on disposable income per capita for the three first cohorts up to 45 months.

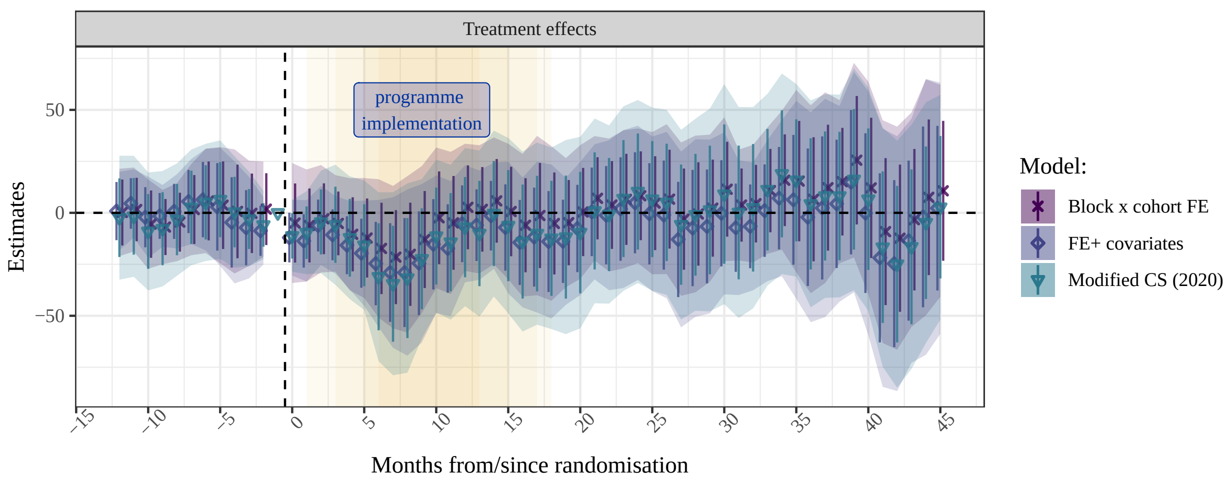
Change in disposable income per consumption units

Cohort 2018 to 2020 only



Intention−to−treat: dynamic difference−in−differences

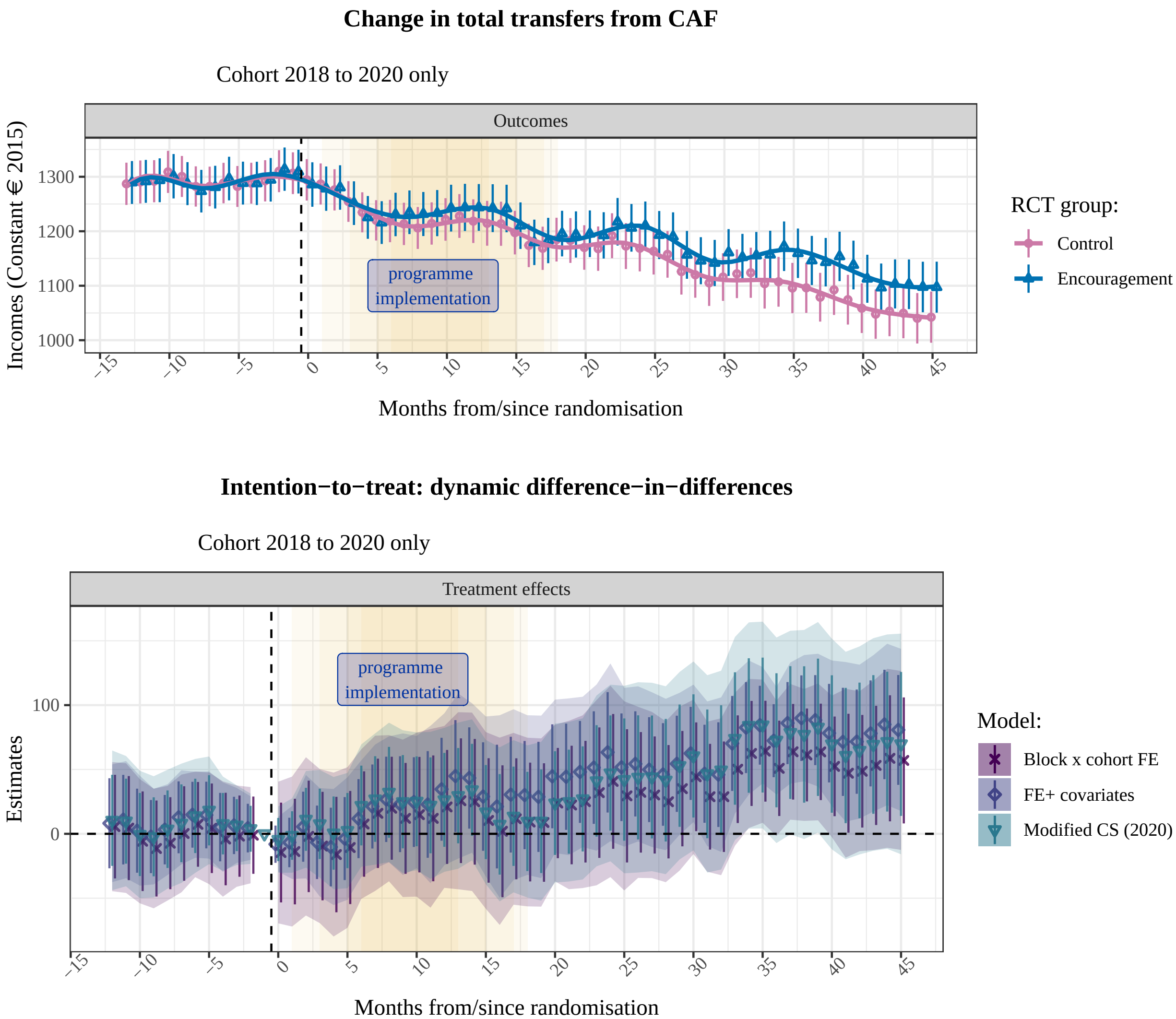
Cohort 2018 to 2020 only



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the monthly disposable income per consumption units, in 2015 euros.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

C) Intention-to-treat for cohort 2018 to 2020 on total transfers

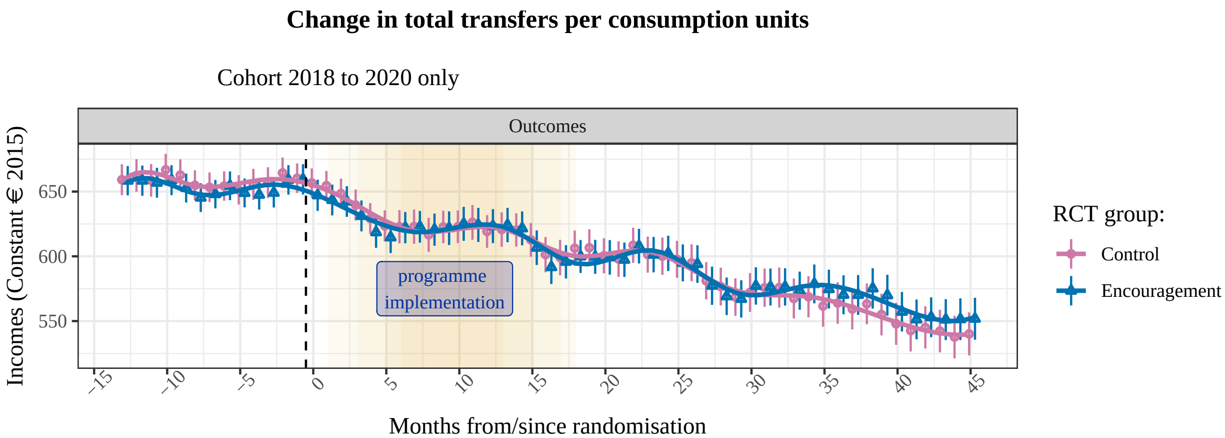
Figure E.34: Intention-to-treat effects of the programme on total cash transfers for the three first cohorts up to 45 months.

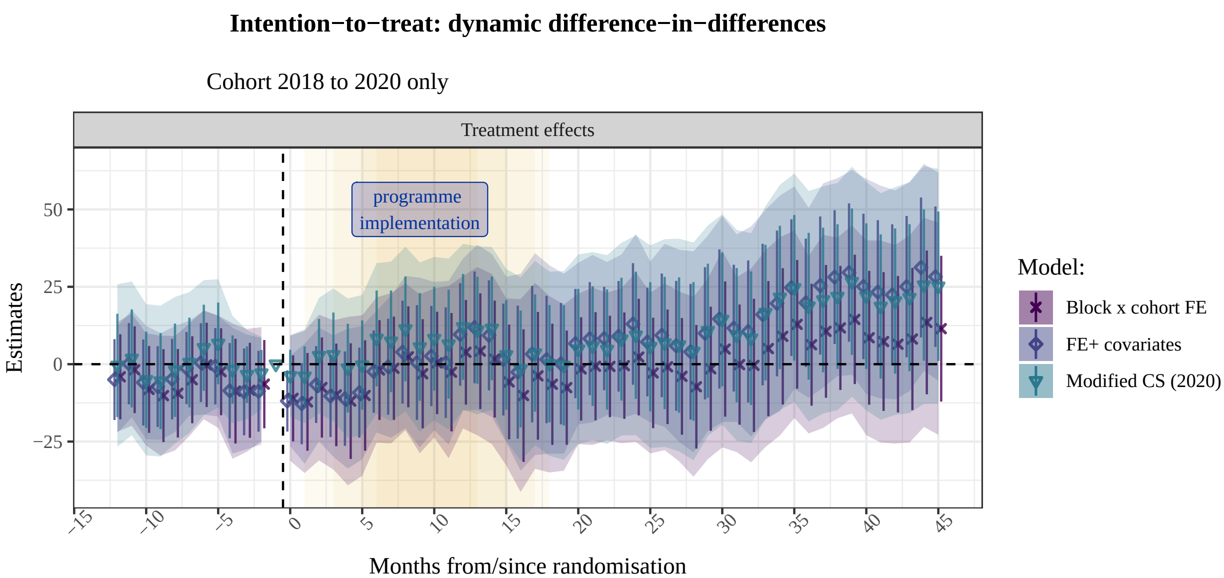


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the monthly total allowance paid by CAF, in 2015 euros.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

D) Intention-to-treat for cohort 2018 to 2020 on total transfers per consumption units

Figure E.35: Intention-to-treat effects of the programme on total cash transfers per consumption units for the three first cohorts up to 45 months.





Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the monthly total allowance paid by CAF in 2015 euros divided by the number of consumption units.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

Bibliography

Abadia, Laura, Ilf Bencheikh, Marie Borel, and Marc Gurgand. 2017. "Accompagnement Des Chômeurs de Longue Durée: Revue Systématique Des Évaluation d'impact." Paris: J-pal.

Abadie, A. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113 (2): 231–63.

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies* 72 (1): 1–19.

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2020. "Sampling-Based Versus Design-Based Uncertainty in Regression Analysis." *Econometrica* 88 (1): 265–96.

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138 (1): 1–35.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Christopher R. Walters. 2018. "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economics* 10 (1): 175–206.

Angrist, J. D., and Guido W Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430).

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation*." *The Quarterly Journal of Economics* 132 (2): 871–919.

Angrist, Joshua, Peter Hull, and Christopher Walters. 2023. "Methods for Measuring School Effectiveness." In *Handbook of the Economics of Education*. Elsevier.

Anne, Denis, Laetitia Challe, Pascale Petit, Sylvain Chareyron, Yannick L'Horty, and Loïc du Parquet. 2019. "Discrimination Dans Le Recrutement Des Grandes Entreprises : Une Approche Multicanal." Rapport d'\'evaluation. Paris: CGET.

Arni, Patrick, Rafael Lalive, and Jan C. Van Ours. 2013. "How Effective Are Unemployment Benefit Sanctions? Looking Beyond Unemployment Exit." *Journal of Applied Econometrics* 28 (7): 1153–78.

Askenazy, Philippe. 2022. "The Transformations of the French Labor Market, 2000–2021." *IZA World of Labor*.

Athey, S., and G. W. Imbens. 2017a. "Chapter 3 - the Econometrics of Randomized Experiments." In *Handbook of Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:73–140. Handbook of Economic Field Experiments. North-Holland.

———. 2017b. "Chapter 3 - The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:73–140. Handbook of Field Experiments. North-Holland.

Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32.

Avram, Silvia, Mike Brewer, and Andrea Salvatori. 2018. "Can't Work or Won't Work: Quasi-experimental Evidence on Work Search Requirements for Single Parents." *Labour Economics* 51 (April): 63–85.

Baronnet, Juliette, Alice Best, Florence Brunet, and Nicolas Duvoux. 2021. "Accompagner Les Familles Monoparentales Moyens Et Enjeux de l'autonomisation de Publics à La Croisée Des Vulnérabilités." Dossier d'\'etude 225. Paris: Cnaf.

Barr, Andrew, Jonathan Eggleston, and Alexander A Smith. 2022. "Investing in Infants: The Lasting Effects of Cash Transfers to New Families*." *The Quarterly Journal of Economics*, April, qjac023.

Behaghel, Luc, Bruno Crépon, and Marc Gurgand. 2014. "Private and Public Provision of Counseling to Job Seekers: Evidence from a Large Controlled Experiment." *American Economic Journal: Applied Economics* 6 (4): 142–74.

Bergemann, Annette, and Gerard J. Van Den Berg. 2008. "Active Labor Market Policy Effects for Women in Europe — A Survey." *Annales d'Économie Et de Statistique*, no. 91/92: 385–408.

Betkó, János. 2023. "Effects of Welfare Policies Based on Autonomy and Unconditionality." PhD thesis, Radboud University.

Blandhol, Christine, Magne Mogstad, John Bonney, and Alexander Torgovitsky. 2022. "When Is TSLS Actually LATE?" Becker Friedman Institute for Economics at UChucago. February 9, 2022.

Bloom, Howard S., Carolyn J. Hill, and James A. Riccio. 2003. "Linking Program Implementation and Effectiveness: Lessons from a Pooled Sample of Welfare-to-Work Experiments." *Journal of Policy Analysis and Management* 22 (4): 551–75.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016. "Female Labor Supply, Human Capital, and Welfare Reform." *Econometrica* 84 (5): 1705–53.

Bono, Pierre-Henri, Rozenn Desplatz, Sandra Debu, and Clément Lacouette-Fougère. 2021. "Le Lent Développement Des Évaluations d'impact En France : Une Approche Par Les Acteurs." *Revue Francaise d'administration Publique* N° 177 (1): 17–28.

Bonoli, Giuliano. 2011. "Active Labour Market Policy and Social Investment: A Changing Relationship," November.

Borusyak, Kirill, UC Berkeley, and Peter Hull. 2023. "Design-Based Identification with Formula Instruments: A Review."

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-Experimental Shift-Share Research Designs." Edited by Dirk Krueger. *The Review of Economic Studies* 89 (1): 181–213.

Brady, Michelle, and Kay Cook. 2015. "The Impact of Welfare to Work on Parents and Their Children." *Evidence Base* 3: 23.

Brodeur, Abel, Nikolai Cook, and Anthony Heyes. 2020. "Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics." *American Economic Review* 110 (11): 3634–60.

Broussard, C. Anne. 2010. "Research Regarding Low-Income Single Mothers' Mental and Physical Health: A Decade in Review." *Journal of Poverty* 14 (4): 443–51.

Brown, Elizabeth. 2020. "Parcours Conjugaux, Violence Conjugale Et Différences de Genre." In *Violences Et Rapports de Genre*, by Elizabeth Brown, Alice Debauche, Christelle Hamel, and Magali Mazuy, 38. Grandes Enquêtes.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2020a. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, December, S0304407620303948.

———. 2020b. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, December, S0304407620303948.

Campbell, Mhairi, Hilary Thomson, Candida Fenton, and Marcia Gibson. 2016. "Lone Parents, Health, Wellbeing and Welfare to Work: A Systematic Review of Qualitative Studies." *BMC Public Health* 16 (1): 188.

Carcasson, Martín. 2006. "Ending Welfare as We Know It: President Clinton and the Rhetorical Transformation of the Anti-Welfare Culture." *Rhetoric and Public Affairs* 9 (4): 655–92.

Card, David, Jochen Kluve, and Andrea Weber. 2010. "Active Labour Market Policy Evaluations: A Meta-Analysis." *The Economic Journal* 120 (548): F452–77.

———. 2018. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* 16 (3): 894–931.

Castell, Laura, Marc Gurgand, Clement Imbert, and Todor Tochev. 2022. "Take-up of Social Benefits: Experimental Evidence from France." December 2022.

Cattaneo, Matias D. 2010. "Efficient Semiparametric Estimation of Multi-Valued Treatment Effects Under Ignorability." *Journal of Econometrics* 155 (2): 138–54.

Cattoen, Victoire, Clotilde Leire, Nathanaël Trellet, and Izaya Youssouf. 2022. "Parcours de Vie Et Expériences Du RSA. Une Approche Sociologique." Université de Lille.

Cervera, Melaine, Céline Émond, Renaud Hourcade, Céline Jung, and Rémi Le Gall. 2017. "L'accompagnement Social Et Ses Effets Sur Les Bénéficiaires." Revue de litt\'erature critique 195. Caisse nationale des allocations familiales.

Chachou, Khadija. 2019. "Le Non-Recours Au Droit à l'accompagnement Des Publics Allocataires Du RSA: Analyse Des Facteurs de La Non-Participation." M\'emoire de Master. Nancy: Université de Lorraine.

Chaisemartin, C De, and Xavier D'Haultfoeuille. 2017. "Fuzzy Differences-in-Differences." *Review of Economic Studies*, no. 1: 1–30.

Chaisemartin, Clément de, and Jaime Ramirez-Cuellar. 2022. "At What Level Should One Cluster Standard Errors in Paired and Small-Strata Experiments?" w27609. Cambridge, MA: National Bureau of Economic Research.

Chan, Marc K, and Robert Moffitt. 2018. "Welfare Reform and the Labor Market." *Annual Review of Economics* 10: 347–81.

Chareyron, Sylvain, David Gray, and Yannick L'Horty. 2018. "Raising Take-up of Social Assistance Benefits Through a Simple Mailing: Evidence from a French Field Experiment." *Revue d'économie Politique* 128 (5): 777–805.

Chareyron, Sylvain, Rémi Le Gall, and Yannick L'Horty. 2022. "Droits Et Devoirs Du RSA : L'impact Des Contrôles Sur La Participation Des Bénéficiaires:" *Revue Économique* Vol. 73 (5): 735–59.

Combe, Julien, and Arthur Heim. 2024. "Rage Against the Matching: Fairness and Inequalities in a Market Design Experiment of Daycare Assignments in France." In *Social Investment & The Changing Face of Poverty: Essays on the Design and Evaluation of Family and Social Policies in France*, Paris School of Economics. 1. Paris, France: PhD Thesis.

Cour des comptes. 2022. "Le Revenu de Solidarité Active." Paris: Cours des comptes.

Crepon, Bruno, Marc Gurgand, Thierry Kamionka, and Laurent Lequien. 2013. "Is Counseling Welfare Recipients Cost-Effective? Lessons from a Random Experiment." Série Des Documents de Travail. CREST. February 2013.

Crépon, Bruno, and Gerard J. van den Berg. 2016. "Active Labor Market Policies." *Annual Review of Economics* 8 (1): 521–46.

Damon, Julien. 2018. "La Juste Prestation : Pour Des Prestations Et Un Accompagnement Ajustés." Rapport au 1er Ministre.

Dardier, Agathe, Quynh-Chi Doan, and Cléo Lhermet. 2022. "La Revalorisation Du Bonus Individuel En 2019 a Fortement Élargi Le Champ Des Bénéficiaires de La Prime d'activité." 1225. Paris: DREES.

DARES. 2021. "Expérimentation Territoire Zéro Chômeurs de Longue Durée." Paris: DARES.

Delahais, Thomas, and Clément Lacouette-Fougère. 2019. "Try Again. Fail Again. Fail Better. Analysis of the Contribution of 65 Evaluations to the Modernisation of Public Action in France." *Evaluation* 25 (2): 131–48.

Delattre, Eric, Richard K. Moussa, and Mareva Sabatier. 2019. "Health Condition and Job Status Interactions: Econometric Evidence of Causality from a French Longitudinal Survey." *Health Economics Review* 9 (1): 3.

Ding, Peng, Avi Feller, and Luke Miratrix. 2019. "Decomposing Treatment Effect Variation." *Journal of the American Statistical Association* 114 (525): 304–17.

DREES. 2022. "Minima Sociaux Et Prestations Sociales Ménages Aux Revenus Modestes Et Redistribution." PANORAMAS DE LA DREES SOCIAL.

Duncan, Greg J., Katherine Magnuson, Ariel Kalil, and Kathleen Ziol-Guest. 2012. "The Importance of Early Childhood Poverty." *Social Indicators Research* 108 (1): 87–98.

Duncan, Greg J., Kathleen M. Ziol-Guest, and Ariel Kalil. 2010. "Early-Childhood Poverty and Adult Attainment, Behavior, and Health." *Child Development* 81 (1): 306–25.

Duriancik, David M, and Courtney R Goff. 2019. "Children of Single-Parent Households Are at a Higher Risk of Obesity: A Systematic Review." *Journal of Child Health Care* 23 (3): 358–69.

Durovic, Anja. 2023. "Rising Electoral Fragmentation and Abstention: The French Elections of 2022." *West European Politics* 46 (3): 614–29.

Duvoux, Nicolas, and Michèle Lelièvre. 2021. "La Pauvreté Démultipliée Dimensions, Processus Et Réponses Printemps 2020 | Printemps 2021." Rapport au premier ministre. Conseil national des politiques de lutte contre la pauvreté et l'exclusion sociale.

Dwyer, Peter, Lisa Scullion, Katy Jones, Jenny McNeill, and Alasdair B. R. Stewart. 2020. "Work, Welfare, and Wellbeing: The Impacts of Welfare Conditionality on People with Mental Health Impairments in the UK." *Social Policy & Administration* 54 (2): 311–26.

Ellwood, David T. 2000. "Anti-Poverty Policy for Families in the Next Century: From Welfare to Work—and Worries." *Journal of Economic Perspectives* 14 (1): 187–98.

Evans, Gary W., and Kalee De France. 2022. "Childhood Poverty and Psychological Well-Being: The Mediating Role of Cumulative Risk Exposure." *Development and Psychopathology* 34 (3): 911–21.

Evans, Patricia M. 2007. "(Not) Taking Account of Precarious Employment: Workfare Policies and Lone Mothers in Ontario and the UK." *Social Policy & Administration* 41 (1): 29–49.

Filges, Trine, G Smedslund, Anne-Sophie Due Knudsen, and Anne-Marie Klint Jørgensen. 2015. "Active Labour Market Programme (ALMP) Participation for Unemployment Insurance Recipients." The Campbell

Collaboration.
Finkelstein, Amy, and Matthew J Notowidigdo. 2019. “Take-Up and Targeting: Experimental Evidence from SNAP*.” *The Quarterly Journal of Economics* 134 (3): 1505–56.
FORS. 2020. “Mission d’évaluation Du Projet d’investissement Social RELIANCE Rapport Final.” Rapport d'\'evaluation. Conseil Départemental Meurthe-Et-Moselle.
Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. “Decomposition Methods in Economics.” In *Handbook of Labor Economics*, 4:1–102. Elsevier.
Foster, Carly Hayden. 2008. “The Welfare Queen: Race, Gender, Class, and Public Opinion.” *Race, Gender & Class* 15 (3/4): 162–79.
Freedman, Seth M., Alex Hollingsworth, Kosali I. Simon, Coady Wing, and Madeline Yozwiak. 2023. “Designing Difference in Difference Studies With Staggered Treatment Adoption: Key Concepts and Practical Guidelines.” Working Paper. Working Paper Series. November 2023.
Friedrichsen, Jana, Tobias König, and Renke Schmacker. 2018. “Social Image Concerns and Welfare Take-up.” *Journal of Public Economics* 168 (December): 174–92.
Frölich, Markus, and Blaise Melly. 2013. “Identification of Treatment Effects on the Treated with One-Sided Non-Compliance.” *Econometric Reviews* 32 (3): 384–414.
Galitzine, Alexandra, and Arthur Heim. 2024. “Tax Burden on the Poor: Single Mothers’ Optimisation Behaviours Following an Experimental Activation Programme in France.” In *Social Investment & The Changing Face of Poverty: Essays on the Design and Evaluation of Family and Social Policies in France*, Paris School of Economics. Paris (France): PhD Thesis.
Gazier, Bernard. 2019. “Opportunities or Tensions: Assessing French Labour Market Reforms from 2012 to 2018.” *International Journal of Comparative Labour Law and Industrial Relations* 35 (September): 331–54.
Gendre, Alexandra de, Stefanie Schurer, and Angela Zhang. 2022. “Two Decades of Welfare Reforms in Australia: How Did They Affect Single Mothers and Their Children?” *SSRN Electronic Journal*.
Gennetian, Lisa, Greg Duncan, Nathan Fox, Katherine Magnuson, Sarah Halpern-Meekin, Kimberly Noble, and Hirokazu Yoshikawa. 2022. “Unconditional Cash and Family Investments in Infants: Evidence from a Large-Scale Cash Transfer Experiment in the U.S.” w30379. Cambridge, MA: National Bureau of Economic Research.
Gibson, Marcia, Hilary Thomson, Kasia Banas, Vittoria Lutje, Martin J McKee, Susan P Martin, Candida Fenton, Clare Bambra, and Lyndal Bond. 2018. “Welfare-to-Work Interventions and Their Effects on the Mental and Physical Health of Lone Parents and Their Children.” Edited by Cochrane Public Health Group. *Cochrane Database of Systematic Reviews* 2018 (2).
Gorey, Kevin. 2009. “Welfare-to-Work Programs in America, 1980 to 2005: Meta-Analytic Evidence of the Importance of Job and Child Care Availability.” *Journal of Policy Practice* 8 (4): 265–81.
Goux, Dominique, and Eric Maurin. 2010. “Public School Availability for Two-Year Olds and Mothers’ Labour Supply.” *Labour Economics* 17 (6): 951–62.
Gurgand, Marc, and David N. Margolis. 2008. “Does Work Pay in France? Monetary Incentives, Hours Constraints, and the Guaranteed Minimum Income.” *Journal of Public Economics* 92 (7): 1669–97.
Hargreaves, Alec G. 2015. “Empty Promises?: Public Policy Against Racial and Ethnic Discrimination in France.” *French Politics, Culture & Society* 33 (3): 95–115.
Heckman, James J., Robert J. Lalonde, and Jeffrey A. Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs.” In *Handbook of Labor Economics*, 3:1865–2097. Elsevier.
Heckman, James J., and Jeffrey A. Smith. 1999. “The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies.” *The Economic Journal* 109 (457): 313–48.
Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. “Characterizing Selection Bias Using Experimental Data.” *Econometrica* 66 (5): 1017–98.
Hemerijck, A. C. 2014. “Social Investment ‘Stocks,’ ‘Flows’ and ‘Buffers’.” *Rivista Delle Politiche Sociali* 1 (1): 9–26.
Herbst-Debby, Anat. 2022. “(De)legitimization of Single Mothers’ Welfare Rights: United States, Britain and Israel.” *Journal of European Social Policy* 32 (3): 302–16.
Hewlett, Nick, and Raymond Kuhn. 2022. “Reflections on the 2022 Elections in France.” *Modern & Contemporary France* 30 (4): 393–409.
Hirano, Kiesukie, Guido W Imbens, and Geert Ridder. 2003. “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score.” *Econometrica* 71 (4): 1161–89.

Hirschman, Daniel, and Elizabeth Popp Berman. 2014. "Do Economists Make Policies? On the Political Effects of Economics1." *Socio-Economic Review* 12 (4): 779–811.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. "Simultaneous Inference in General Parametric Models." *Biometrical Journal* 50 (3): 346–63.

INSEE. 2023. "Dossier Complet - Arrondissement de Nancy (543) | Insee." January 2023.

Jacquey-Vazquez, Benedicte. 2017. "L'accompagnement Social." RAPPORT IGAS 2017-105K.

Jaehrling, K., T. Kalina, and L. Mesaros. 2015. "A Paradox of Activation Strategies: Why Increasing Labour Market Participation Among Single Mothers Failed to Bring Down Poverty Rates." *Social Politics: International Studies in Gender, State & Society* 22 (1): 86–110.

Jaroszewicz, Ania, Jon Jachimowicz, Oliver Hauser, and Julian Jamison. 2022. "How Effective Is (More) Money? Randomizing Unconditional Cash Transfer Amounts in the US." *SSRN Electronic Journal*.

Jenson, Jane. 2010. "Diffusing Ideas for After Neoliberalism: The Social Investment Perspective in Europe and Latin America." *Global Social Policy: An Interdisciplinary Journal of Public Policy and Social Development* 10 (1): 59–84.

Johnsen, Julian Vedeler, and Katrine Holm Reiso. 2020. "Economic Effects of Workfare Reforms for Single Mothers: Benefit Substitution and Labour Supply Responses*." *The Scandinavian Journal of Economics* 122 (2): 494–523.

Jones, Damon, and Ioana Marinescu. 2022. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." *American Economic Journal: Economic Policy* 14 (2): 315–40.

Kasy, Maximilian, and Lukas Lehner. 2023. "Employing the Unemployed of Marienthal: Evaluation of a Guaranteed Job Program."

Kerr, N. L. 1998. "HARKing: Hypothesizing After the Results Are Known." *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc* 2 (3): 196–217.

Knapp, Andrew. 2022. "France's Party System in 2022." *Modern & Contemporary France* 30 (4): 495–515.

Knijn, Trudie, Claude Martin, and Jane Millar. 2007. "Activation as a Common Framework for Social Policies Towards Lone Parents." *Social Policy & Administration* 41 (6): 638–52.

Ko, Wonsik, and Robert A Moffitt. 2022. "Take-up of Social Benefits." Discussion Paper. 2022.

Kramarz, Francis, Elio Nimier-David, and Thomas Delemotte. 2022. "Inequality and Earnings Dynamics in France: National Policies and Local Consequences." *Quantitative Economics* 13 (4): 1527–91.

Lalonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *The American Economic Review* 76 (4): 604–20.

Latzman, Natasha E., Cecilia Casanueva, Julia Brinton, and Valerie L. Forman-Hoffman. 2019. "The Promotion of Well-Being Among Children Exposed to Intimate Partner Violence: A Systematic Review of Interventions." *Campbell Systematic Reviews* 15 (3): e1049.

Le Pape, Marie Clémence, and Clémence Helfter, eds. 2023. *Les Familles Monoparentales: Condition de Vie, Vécu Et Action Publique*. États Des Savoirs. Paris: La Documentation française.

Leruth, Benjamin. 2017. "France at a Crossroads: Societal Challenges to the Welfare State During Nicolas Sarkozy's and François Hollande's Presidential Terms." In *After Austerity: Welfare State Transformation in Europe After the Great Recession*, edited by Peter Taylor-Gooby, Benjamin Leruth, and Heejung Chung, 0. Oxford University Press.

Liang, Kung-Yee, and Scott L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73 (1): 13–22.

Lin, Winston. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *The Annals of Applied Statistics* 7 (1): 295–318.

Løken, Katrine V., Kjell Erik Lommerud, and Katrine Holm Reiso. 2018. "Single Mothers and Their Children: Evaluating a Work-Encouraging Welfare Reform." *Journal of Public Economics* 167 (November): 1–20.

Lucas, Patricia J., Karen McIntosh, Mark Petticrew, Helen M. Roberts, and Alan Shiell. 2008. "Financial Benefits for Child Health and Well-Being in Low Income or Socially Disadvantaged Families in Developed World Countries." *Campbell Systematic Reviews* 4 (1): 1–93.

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. 2023. "Cluster-Robust Inference: A Guide to Empirical Practice." *Journal of Econometrics* 232 (2): 272–99.

Mahdi, Samira. 2021. "Innovation Sociale: Évaluation Des Effets d'un Dispositif Innovant Sur Les Professionnels de l'accompagnement Socioprofessionnel: Le Cas Du Projet Reliance Sur Le Département de Meurthe Et-Moselle." M\'emoire de Master Sciences de l'\'education « Formation, travail, Territoires et d\'eveloppement

». Nancy: Université de Lorraine.

Mangin, Evelyn Jean. 2021. "A Narrative Policy Framework and Feminist Critical Policy Analysis of the Welfare Reform Acts, and The Higher Education of Single Welfare Mothers." PhD thesis, Walden University.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149): 976–80.

Markussen, Simen, and Knut Røed. 2016. "Leaving Poverty Behind? The Effects of Generous Income Support Paired with Activation." *American Economic Journal: Economic Policy* 8 (1): 180–211.

Martin, John P. 2015. "Activation and Active Labour Market Policies in OECD Countries: Stylised Facts and Evidence on Their Effectiveness." *IZA Journal of Labor Policy* 4 (1): 4.

McCRATE, ELAINE, and JOAN SMITH. 1998. "WHEN WORK DOESN'T WORK: The Failure of Current Welfare Reform." *Gender & Society* 12 (1): 61–80.

McLanahan, Sara. 2004. "Diverging Destinies: How Children Are Faring Under the Second Demographic Transition." *Demography* 41 (4): 607–27.

Meisel, Joan, Daniel Chandler, and Beth Menees Rienzi. 2003. "Domestic Violence Prevalence and Effects on Employment in Two California TANF Populations." *Violence Against Women* 9 (10): 1191–1212.

Millar, Michelle, Jane Gray, Clíona Rooney, and Rosemary Crosse. 2018. "The (in)adequacy of in-Work Benefits in Irish Lone Parent Labour Market Activation Policy." *Journal of Poverty and Social Justice* 26 (3): 379–400.

Milner, Susan. 2017. "Employment and Labour Market Policy During the Hollande Presidency: A Tragedy in Three Acts?" *Modern & Contemporary France* 25 (4): 429–43.

Moffitt, Robert A. 2015. "The Deserving Poor, the Family, and the U.S. Welfare System." *Demography* 52 (3): 729–49.

Mogstad, Magne, and Chiara Pronzato. 2012. "Are Lone Mothers Responsive to Policy Changes? Evidence from a Workfare Reform in a Generous Welfare State." *The Scandinavian Journal of Economics* 114 (4): 1129–59.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters. 2021. "The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables." *American Economic Review* 111 (11): 3663–98.

Moller, Stephanie. 2002. "Supporting Poor Single Mothers: Gender and Race in the U.S. Welfare State." *Gender and Society* 16 (4): 465–84.

Morel, Nathalie, Bruno Palier, and Joakim Palme. 2012. *Towards a Social Investment Welfare State? : Ideas, Policies and Challenges*. Edited by Nathalie Morel, Bruno Palier, and Joakim Palme. Policy Press.

Morescalchi, Andrea, and Paolo Paruolo. 2020. "Too Much Stick for the Carrot? Job Search Requirements and Search Behaviour of Unemployment Benefit Claimants." *The B.E. Journal of Economic Analysis & Policy* 20 (1).

Negi, Akanksha, and Jeffrey M. Wooldridge. 2021. "Revisiting Regression Adjustment in Experiments with Heterogeneous Treatment Effects." *Econometric Reviews* 40 (5): 504–34.

Nieuwenhuis, Rense. 2020. "The Situation of Single Parents in the EU." European parliament.

Pearl, Judea. 2009. *Causality Models, Reasoning and Inference*. Cambridge, MA: Cambridge University Press.

Peden, George. 2017. "Liberal Economists and the British Welfare State." In *Liberalism and the Welfare State: Economists and Arguments for the Welfare State*, Roger E. Backhouse, Bradley W. Bateman, Tamotsu Nishizawa and Dieter Plehwe, 1:39–56. Oxford University Press.

Perez, Sara, Dawn M. Johnson, and Caroline Vaile Wright. 2012. "The Attenuating Effect of Empowerment on IPV-Related PTSD Symptoms in Battered Women Living in Domestic Violence Shelters." *Violence Against Women* 18 (1): 102–17.

Périvier, Hélène, and Muriel Pucci. 2019. "Le Recouvrement Des Impayés de Pensions Alimentaires Réduit Les Dépenses Sociales Mais Réduit Également Le Niveau de Vie de Certaines Mères Isolées." *Sciences Po Publications*, no. info:hdl:2441/2d6dp3s2nh8dmpd40eh37dd6pe (May).

Périvier, Hélène, and Réjane Sénac. 2017. "The New Spirit of Neoliberalism: Equality and Economic Prosperity." *International Social Science Journal* 67 (223-224): 31–42.

Perrin-Heredia, Ana. 2009. "Les Logiques Sociales de l'endettement : Gestion Des Comptes Domestiques En Milieux Populaires." *Sociétés Contemporaines* 76 (4): 95–119.

Peterson, Janice. 1997. ""Ending Welfare as We Know It": The Symbolic Importance of Welfare Policy in America." *Journal of Economic Issues* 31 (2): 425–31.

Pitollat, Claire, and Matthieu Klein. 2018, Aout. "L'accompagnement Des Bénéficiaires Du RSA." Rapport au 1er Ministre. Paris.

Redman, Jamie. 2020. "The Benefit Sanction: A Correctional Device or a Weapon of Disgust?" *Sociological*

Research Online 25 (1): 84–100.
Reese, Ellen. 2005. *Backlash Against Welfare Mothers: Past and Present*. 1st ed. University of California Press.
Reingold, Beth, and Adrienne R. Smith. 2012. "Welfare Policymaking and Intersections of Race, Ethnicity, and Gender in U.S. State Legislatures." *American Journal of Political Science* 56 (1): 131–47.
Riger, Stephanie, and Maryann Krieglstein. 2000. "The Impact of Welfare Reform on Men's Violence Against Women." *American Journal of Community Psychology* 28 (5): 631–47.
Rinaldi, Chiara, and Marleen P. M. Bekker. 2021. "A Scoping Review of Populist Radical Right Parties' Influence on Welfare Policy and Its Implications for Population Health in Europe." *International Journal of Health Policy and Management* 10 (3): 141–51.
Rodríguez, Jorge. 2023. "Understanding the Effects of Workfare Policies on Child Human Capital." *Journal of Labor Economics* 41 (1): 39–75.
Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.
Roth, Jonathan, Pedro H C Sant'Anna, Alyssa Bilinski, and John Poe. 2021. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," 54.
Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics* 235 (2): 2218–44.
Sant'Anna, Pedro H. C., and Jun Zhao. 2020. "Doubly Robust Difference-in-Differences Estimators." *Journal of Econometrics* 219 (1): 101–22.
Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan. 2016. "The Psychological Lives of the Poor." *American Economic Review* 106 (5): 435–40.
Shahidi, Faraz V, Chantel Ramraj, Odmaa Sod-Erdene, Vincent Hildebrand, and Arjumand Siddiqi. 2019. "The Impact of Social Assistance Programs on Population Health: A Systematic Review of Research in High-Income Countries." *BMC Public Health* 19 (1): 2.
Słoczyński, Tymon, S. Derya Uysal, and Jeffrey M. Wooldridge. 2022. "Doubly Robust Estimation of Local Average Treatment Effects Using Inverse Probability Weighted Regression Adjustment." August 2, 2022.
Smedslund, G. 2006. "Work Programmes for Welfare Recipients." *Campbell Systematic Reviews*, August.
Spencer, Rachael A., Melvin D. Livingston, Briana Woods-Jaeger, Shelby T. Rentmeester, Nolan Sroczynski, and Kelli A. Komro. 2020. "The Impact of Temporary Assistance for Needy Families, Minimum Wage, and Earned Income Tax Credit on Women's Well-Being and Intimate Partner Violence Victimization." *Social Science & Medicine (1982)* 266 (December): 113355.
Sun, Liyang, and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, December, S030440762030378X.
Tan, Zhiqiang. 2010. "Bounded, Efficient and Doubly Robust Estimation with Inverse Weighting." *Biometrika* 97 (3): 661–82.
Tübbicke, S. 2023. "When to Use Matching and Weighting or Regression in Instrumental Variable Estimation? Evidence from College Proximity and Returns to College." *Empirical Economics* 65 (6): 2979–99.
Tur-Prats, Ana. 2019. "Family Types and Intimate Partner Violence: A Historical Perspective." *The Review of Economics and Statistics* 101 (5): 878–91.
Van Winkle, Zachary, and Emanuela Struffolino. 2018. "When Working Isn't Enough: Family Demographic Processes and in-Work Poverty Across the Life Course in the United States." *Demographic Research* 39 (September): 365–80.
Vandenbroucke, Frank, and Koen Vleminckx. 2011. "Disappointing Poverty Trends: Is the Social Investment State to Blame?" *Journal of European Social Policy* 21 (5): 450–71.
Verlaat, Timo, Stephanie Rosenkranz, Loek F. M. Groot, and Mark Sanders. 2021. "Requirements Vs. Autonomy: What Works in Social Assistance?" SSRN Scholarly Paper. Rochester, NY. June 21, 2021.
Vooren, Melvin, Carla Haelermans, Wim Groot, and Henriëtte Maassen van den Brink. 2019. "The Effectiveness of Active Labor Market Policies: A Meta-Analysis." *Journal of Economic Surveys* 33 (1): 125–49.
Wunsch, Conny. 2016. "How to Minimize Lock-in Effects of Programs for Unemployed Workers." *IZA World of Labor*, January.
Wüthrich, Kaspar, and Ying Zhu. 2023. "Omitted Variable Bias of Lasso-Based Inference Methods: A Finite Sample Analysis." *The Review of Economics and Statistics* 105 (4): 982–97.