



Pattaro, S. , Bailey, N. , Williams, E. , Gibson, M., Wells, V., Wright, S. , Tranmer, M. and Dibben, C. (2019) Labour Market and Wider Impacts of Benefit Sanctions: A Scoping Review. 17th Annual ESPAnet Conference, Stockholm, Sweden, 5-7 Sept 2019.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/208879/>

Deposited on 27 January 2020

Enlighten – Research publications by members of the University of
Glasgow

<http://eprints.gla.ac.uk>

Labour market and wider impacts of benefit sanctions: A scoping review

Serena Pattaro^{1,2}, Nick Bailey^{1,2}, Evan Williams², Marcia Gibson³, Valerie Wells³,
Sharon Wright², Mark Tranmer², and Chris Dibben^{1,4}

¹ Scottish Centre for Administrative Data Research, University of Edinburgh, Edinburgh, UK.

² School of Social and Political Sciences, University of Glasgow, Glasgow, UK.

³ MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Glasgow, UK.

⁴ School of Geosciences, University of Edinburgh, Edinburgh, UK.

August 2019

[This is a preliminary draft. Please do not quote without permission]

Abstract

Motivation. In recent decades, the use of conditionality backed by benefit sanctions for those claiming unemployment and related benefits have become widespread in the welfare systems of advanced societies. Governments have increased the intensity and scope of sanctions with the aim of encouraging individuals to move off benefits and return to work. Existing reviews assessing the effects of sanctions on benefit exits and labour market outcomes have found some positive impacts although also variations between studies. Evidence from qualitative research has drawn attention to a range of negative consequences for individual health and other social outcomes. To our knowledge, the quantitative studies on such wider impacts have not been rigorously reviewed.

Aims and methods. We conduct a scoping review of the existing international evidence on labour market and wider impacts of benefit sanctions, in order to systematically assess the state of the quantitative literature on this topic. We develop a search strategy based on an extended list of terms and synonyms for benefit sanctions. We combine a search of major bibliographic databases used across the social and health sciences with a hand search of key websites of relevant research and policy organisations. We follow a review protocol to extract the information on the main features of each study, including outcome measures used, impacts identified and characteristics of target populations. We also examine key features of study design and methodologies applied.

Results. We find a total of 109 studies providing original quantitative evidence on the labour market and/or wider impacts of sanctions which meet our other selection criteria (time, language, country). Studies from the US make up the largest group by some way, followed by Western and Northern European studies. Almost two thirds examined labour market outcomes and just one third examined wider impacts; some cover both. While the number of studies is slightly lower in recent years, an increasing proportion are using experimental or quasi-experimental study designs with stronger claims to identify causal effects. Of the studies examining wider impacts, relatively few employ quasi-/experimental designs. The picture emerging from the evidence base appears rather fragmented, with inconsistent findings reported across the study design types, but this maybe due to the limited evidence on the wider impacts.

Conclusions. The quantitative evidence on the wider impacts of benefit sanctions is much thinner than that on labour market impacts. It is further diluted by being spread across a diverse range of outcomes covering adults and their families making it very difficult to draw general conclusions in relation to any particular aspect. Methodologically, the literature on the wider impacts is dominated by studies with weak study designs. It is difficult to conclude that the great bulk of them provide evidence of the causal impact of sanctions and not, for example, the impact of unobserved confounding factors.

Keywords: Benefit sanctions, Social security benefits, Welfare benefits, Unemployment, Scoping review.

1. Introduction

In recent decades, as part of an increasing shift towards active labour market and social policies, the use of conditionality backed by benefit sanctions have become widespread in the welfare systems of advanced societies (Bonoli 2010; Bonoli and Natali 2012).

Governments have increased both the intensity and scope of welfare benefit sanctions with the aim of encouraging working-age individuals to move off unemployment benefits and return to work. Entailing a temporary reduction or interruption of benefit payments, sanctions are imposed on claimants who fail to meet specific conditions related to job search or work preparation (Griggs and Evans 2010). While initially aimed at people unemployed, more recently benefit sanctions have been extended to cover a wider range of population groups, including some of those inactive and/or with long-term sickness or disability and even, in the UK, those in employment (Baumberg Geiger 2017, Dwyer and Wright 2014).

Existing reviews of quantitative studies assessing the effects of sanctions have focussed on the labour market side: on benefit exits and returns to employment. While reported outcomes vary between studies, the evidence suggests that sanctions raise benefit exit rates and (somewhat weaker) hasten returns to employment, but it also suggests that job quality is often poor, that earnings may be lower and that returns to benefits (recidivism) may be higher.

Alongside this, there is a body of largely qualitative research which has drawn attention to a range of possible negative consequences of sanctions on a wide range of areas of life. In one review, Griggs and Evans (2010) highlight studies examining impacts on claimant financial stress, health, propensity to commit crime, and homelessness. In addition, there are impacts on household and family relations, including on child development, education and welfare.

To our knowledge, there has not been a review which has specifically sought to identify quantitative studies on wider impacts. Quantitative studies have two key advantages: first they can provide some measure of the scale of any impacts and second, with the right research design, they can provide evidence that any relationship is caused by the sanction event and not by any other factors.

The present study therefore aims to systematically search for, identify and extract data from the existing international quantitative literature on the labour market and wider impacts of

benefits sanctions. The working-age population in receipt of unemployment-related or other means-tested benefits is the primary focus for this study. We do not apply any restrictions on the outcomes studied as the main purpose is to offer a comprehensive review of the outcomes reported by relevant studies.

We do this by conducting a scoping review, broadly in line with the guidelines developed by Tricco et al. (2018). A scoping review is intended to capture relevant studies on a topic using a systematic search strategy, and study selection and data extraction process. It provides an overview of the nature of the evidence base capturing characteristics such as temporal and/or geographic spread, nature of research methods or study designs and findings. A particular focus here is to establish the nature of the evidence base for wider impacts compared with that for labour market outcomes (relative scale, geographic coverage, outcomes assessed) and to examine the quality of that evidence.

1.1 Research questions

We therefore aim to address the following research questions:

- What is the nature of the evidence base on the impacts of benefit sanctions, for both labour market and wider outcomes?
- What study designs have been used in studies of benefit sanctions, and how do these differ between labour market and wider outcome studies?
- What is the evidence for the wider impacts of sanctions and how does this evidence vary between studies with an experimental or quasi-experimental design and others?

2. Methods

2.1 Scoping review

We draw on the seminal framework by Arksey and O'Malley (2005) and more recent advances (Levac, Colquhoun, and O'Brien 2010; Peters et al. 2015) to identify and systematically synthesise the international evidence from quantitative studies on labour market and the wider impacts of benefit sanctions. We developed a protocol for our scoping review study (Pattaro et al. 2019) by following, where possible, the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines for scoping reviews (Tricco et al. 2018). These are designed to ensure that a rigorous, consistent and transparent process is followed. Scoping reviews belong to the broader family of systematic reviews and aim to answer broad questions and map the existing body of the literature on a particular topic. While valuable in themselves, scoping reviews are also used to inform subsequent systematic reviews, as they provide the baseline knowledge which enables researchers to establish the need to conduct a full systematic review and meta-analysis on a specific research question.

2.2 Search strategy

We iteratively developed an extensive search strategy which encompassed a long list of subject headings, keywords, terms and synonyms for benefit sanctions. The search strategy was developed in consultation with an information officer who is member of the research team and has expertise in systematic reviews (VW). Between March and June 2019, we

conducted initial electronic searches of eight major bibliographic databases used across the social and health sciences: ASSIA, British Education Index, EconLit, ERIC, PsycINFO, MEDLINE, Scopus, SocINDEX. Results are summarised in Table A1 with full details of the search strategies provided in Table A2 (both in the Appendix). In addition, we conducted hand searches of key websites of relevant research and policy organisations (e.g. IZA, NBER, RePEc, IFAU, OECD, ILO). The combined results of the searches were imported into Endnote and deduplicated.

2.3 Inclusion criteria and study selection process

The studies for this review were selected using five inclusion criteria:

- (a) Targeting working-age recipients of welfare- or unemployment-related benefits in high-income countries;
- (b) Investigating sanctions applied to these benefits for failure to comply with work search or other requirements;
- (c) Quantitative studies based on either experimental, quasi-experimental or non-experimental designs;
- (d) In English language;
- (e) Published between January 1990 and February 2019.

The first four authors (SP, NB, EW and MG) conducted the screening and data extraction of the studies included in this review. An overview of the study selection process is shown in Figure 1. Searching across the eight electronic databases yielded 9629 records. These were combined with 401 records retrieved from additional website searches of key research and policy organisations. From the combined databases, 2460 (25%) studies were removed because they were duplicates, leaving a total of 7570 studies.

We conducted initial screening based on the assessment of title and abstract to determine whether studies appeared to meet our eligibility criteria. This led to exclusion of 6387 (84%) studies, because for example the topic of the study was not relevant, their publication date was prior January 1990, or they were not published in English. To ensure the reliability of initial screening, a preliminary review was conducted on 200 studies. The disagreement rate was of 4.5% ($n = 9$) and discrepancies were solved without resorting to a third-party opinion.

Initial screening therefore yielded a sample of 1183 full-text articles that were further assessed during a second screening. This led to the exclusion of a further 851 studies. The majority of these ($n = 596$; 70%) comprised studies whose focus was not on sanctions. In this group there were also studies examining welfare leavers' outcomes and the effects of other welfare reforms such as time limit policies, or job-search interventions not directly reporting sanction impacts. Working papers that were subsequently published as a journal article which was included in our database were also excluded. An additional 197 studies (23%) were excluded due to characteristics pertaining the study design. These included narrative papers based on policy analysis, commentaries, discussion pieces, general overviews, studies based on qualitative analysis, theoretical studies and studies based on

microsimulation modelling. The remaining excluded studies (7%) comprised 27 full-text articles that could not be accessed, 21 out-of-scope studies because they were either published before January 1990, not in English language or not pertaining to high income countries, and 10 studies identified as duplicates at this stage.

This left a sample of 332 studies which were retained for subsequent data extraction and analysis. In this stage of the process, we identified 109 studies (33%) where the authors provided original evidence on the impact of benefit sanctions. The analyses in this paper focus on this analytical subsample of studies.

Of the remainder, 147 studies (44%) were found to be based on the assessment of multiple simultaneous interventions or policy tools so they did not allow the separate identification of the impact of benefit sanctions. These studies were using for example period or policy dummy indicators to identify a set of welfare changes or were combining sanctioned individuals with groups affected by other policies. These were omitted from further analysis.

A further 76 studies (23%) reviewed a number of individual studies without providing original primary evidence themselves. They used a variety of methodologies from more informal narrative reviews to more systematic reviews. These studies were also omitted from further analysis at this stage, on the basis that many of the studies they reviewed should be captured by our database. At a later stage, we will make a comparison of the studies they analysed to see if there are any we can add to our scoping review. Separately, it would be interesting to conduct a review of these reviews, to compare the findings to our own conclusions.

2.4 Data extraction

A data extraction form was developed to record detailed information from the analytical subsample of 109 studies. The form was pilot-tested on a randomly selected study and subsequently finalised on a larger number of studies to ensure it captured all relevant information. The data extraction was carried out in two stages. First, we conducted a partial extraction by gathering information on the main characteristics for all the 109 studies included in the analytical sample. We then focussed on conducting a full extraction of data for those studies reporting wider, non-labour market outcomes. The reasoning behind this restriction lies on the fact that there is quite a large body of literature (e.g. Card, Kluge, and Weber 2010; 2018; Vooren et al. 2019) which focuses on studies reporting labour market and economic outcomes. Little is known on the quality of the evidence base derived from those studies which report non-economic outcomes of benefit sanctions. On these, we gathered information on study design, data sources and sample size, key outcome and exposure measures and main findings.

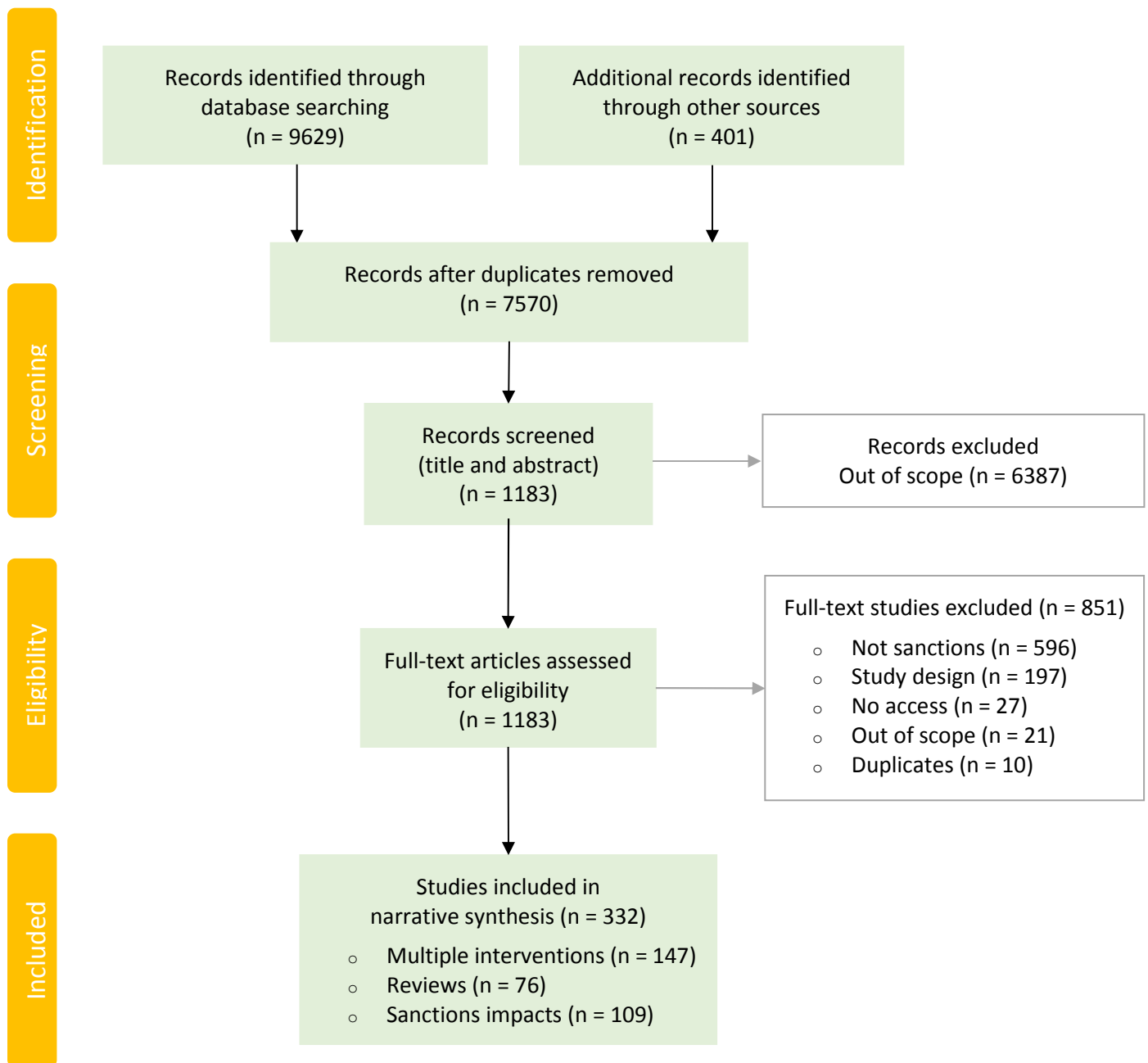


Fig. 1. PRISMA flow chart representing study selection process

2.5 Literature analysis and synthesis

We conduct a descriptive analysis of the evidence base by exploring how this varies by main study characteristics. We use this information to identify relevant patterns in the data and inform the development of a typology based on the main characteristics of study designs. The results emerging from the narrative synthesis based on wider (non-labour market) outcomes are then used to reflect on the extent to which study designs are able to support causal inference.

We start with a three-fold typology for study designs based around their ability to support causal inferences (Murnane and Willett 2010; Angrist and Pischke 2009). Experimental studies have the strongest claims since the researchers work to ensure random allocation to intervention and control groups. This includes randomised controlled trials and studies based on randomised assignment. Quasi-experimental designs exploit exogenous variation occurring 'naturally' through the ways in which policy changes have been introduced or implemented. This group includes difference-in-differences models, regression discontinuity designs and instrumental variables estimation.

Lastly, there are studies with non-experimental designs which rely, to greater or lesser extent, on controlling for differences between sanctioned and non-sanctioned groups using observed characteristics which might also influence the outcome of interest. These include descriptive studies, simple regression models (Ordinary Least Squares (OLS) or logistic regression models), along with more advanced regression models such as survival or time-to-event models, hierarchical or multilevel models, random effects and fixed effects models, and time series models. Propensity-score matching and related approaches also belongs to this group, as they rely on selection on observables to estimate a propensity score, that is the probability of an individual being assigned to an intervention.

3. Results

3.1 Descriptive analysis

Figure 2 presents the overall distribution of studies reporting both labour market and wider outcomes by country and study design ($n = 109$). By far the largest group, which includes 68 studies, is for the USA although many of these cover specific locations within the US or a group of states, rather than the entire country. When combined with the other non-European English-speaking countries (Canada and Australia), these account 65% of the overall sample. The quality of the evidence base, however, is relatively low with a large proportion of studies in the non-experimental group.

The remaining studies are all European and here the quality of study designs tends to be higher. Germany has the highest number of studies ($n = 11$), followed by other Western European countries, such as the Netherlands ($n = 8$), the United Kingdom ($n = 5$) and Switzerland ($n = 4$). Countries from Western Europe (together with Belgium) cover a total of 30 studies, accounting for more than a quarter (27%) of the sample. These are followed by Northern European countries (Denmark, Norway, Finland and Sweden) which cover a total of 8 studies (7% of the sample).

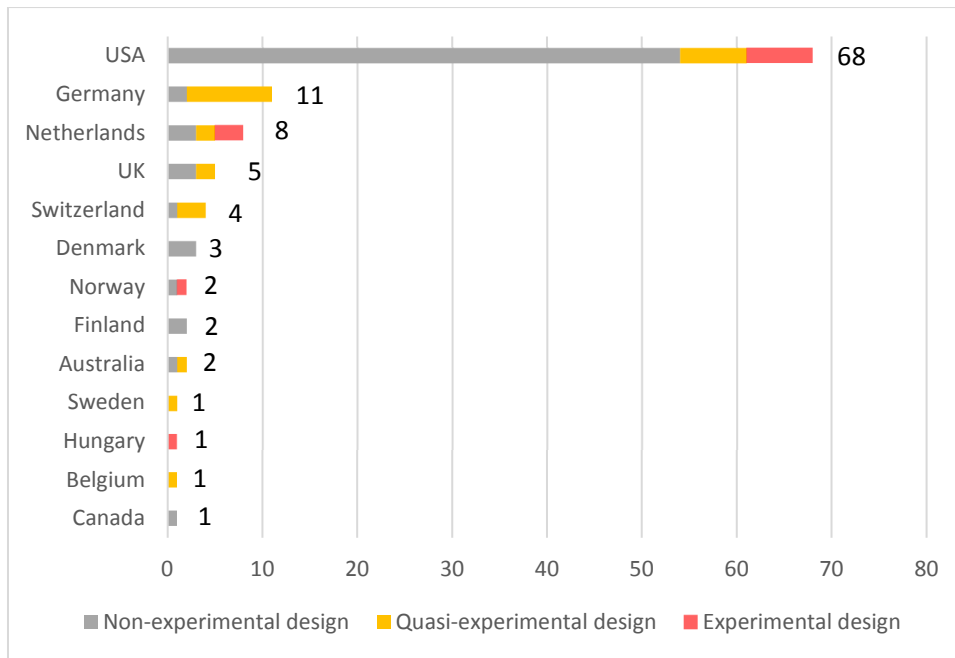


Fig. 2. Number of studies by country and study design

Our sample covers studies published from the mid-1990s onwards (Figure 3 and Table 1). While the number of studies has been on a downward trend, the quality of studies is increasing over time, on average. While experimental designs are sparsely distributed over time, it is clear that there is an increasing trend in the use of quasi-experimental designs starting from the mid-2000s. Within the field of policy intervention evaluations, experimental designs such as randomised controlled trials, although regarded as the ‘gold standard’, require a great amount of resources both in terms of funding and participant involvement.

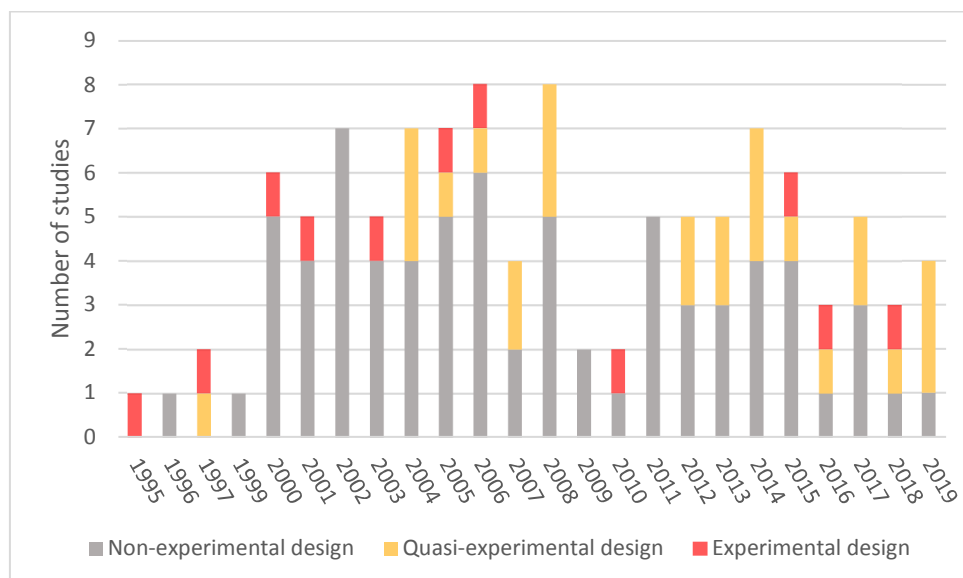


Fig. 3. Distribution of studies by year of publication and study design

Table 1 Frequencies and proportion of studies by period of publication and study design

Publication period	Non-experimental design		Quasi-experimental design		Experimental design		Total (n)
	n	%	n	%	n	%	
1995-1999	2	3	1	4	2	17	5
2000-2009	43	60	10	38	6	50	59
2010-2019	26	37	15	58	4	33	45
	71	100	26	100	12	100	109

On the other hand, quasi-experimental designs are more accessible in the context of evaluation research of policy and social interventions as they can rely more easily on observational data derived from survey and administrative data sources.

Figure 4 shows the distribution of publications arranged by study design and the population group targeted by sanction policy interventions. A large portion (n = 56, 79%) of the studies using an observational design are based on interventions addressing low-income families and lone parents. In terms of both absolute numbers and percentages, the number of studies investigating low-income families or lone parents decreases both in absolute and percentage terms for quasi-experimental (n = 12; 46%) and experimental designs (n = 5; 42%). Conversely, studies based on the unemployed segments of the population, tend to rely more on quasi-experimental (54%) and experimental study designs (58%).

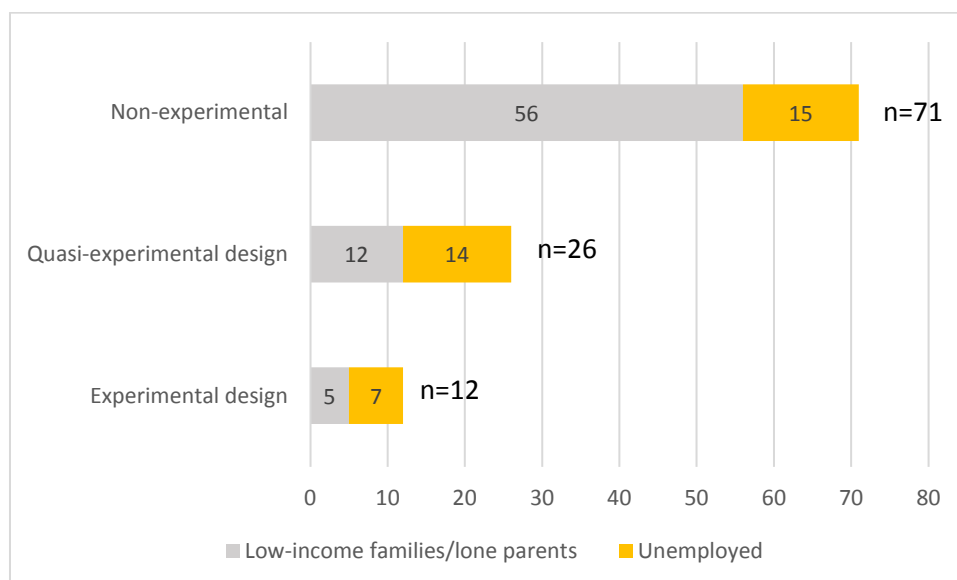


Fig. 4. Number of studies by study design and target population

3.2 Developing a study design typology

So far, we have considered a conventional three-way classification of study designs (non-experimental, quasi-experimental and experimental designs). We refine this classification and develop a study design typology which emerged from the assessment of the studies captured in our analytical sample. Table 2 provides a brief description for each cluster or group and highlights some issues related to the identification of causal effects. The constructed typology is articulated into five types which include a range of estimation tools or methods which are ordered approximately according to the extent to which each method is suitable to address causation.

Types 1 and 2 are non-experimental study designs. Type 1 comprises descriptive studies based on bivariate analysis and studies based on simple multivariable regression techniques, such as linear regression and logistic/probit models. These generally rely on covariate adjustment to account for confounders. Conventional linear and logistic regression analysis lie at the lower end of the continuum generally as they do not provide tools able to support the identification of causal effects. Type 2 includes more advanced regression-based approaches which, depending on how these are implemented, can claim to control for some unmeasured confounding. These include survival and hierarchical models, time series and fixed effects models.

Types 3 and 4 are sometimes bundled together into the quasi-experimental design cluster, although we prefer to limit this designation to Type 4. There is an increasing recognition across the health and social sciences that this design group are characterised by a high degree of heterogeneity in terms of quality of approach and the ability to address causation (Craig et al. 2017; Dunning, 2012). Type 3 includes designs based on matching techniques which rely on covariate adjustment to estimate a propensity score, that is the probability of an individual being assigned to an intervention. Type 4 encompasses difference-in-differences models, regression discontinuity and instrumental variables models which rely on an identification strategy based on exogenous variations occurring 'naturally' in the observed data which allow to support causal claims more effectively.

While propensity score matching and related approaches may still be affected by potential issues arising from both residual and unmeasured confounding, the underlying assumptions of the estimation approaches included in Type 4 may prove difficult to be tested which may raise challenges in the interpretation and attribution of effects. At the highest end of the continuum lie Type 5 which include randomised controlled trials which, by relying on random assignment to allocate individuals to treatment and control groups, are able to minimise sources of selection bias.

Table 2 Overview of the study design typology based on the studies included in the scoping review

Study design typology	Description	Issues for identification of causal effects
1. Descriptive analysis, OLS, logistic/probit regression	Bivariate analyses and multivariable regression models relying on standard covariate adjustment to control for potential confounders	Omission of unobserved confounders which correlate with sanction risks and relevant outcomes may bias estimations of sanction effects.
2. Survival models, hierarchical models, time series, fixed effects	More complex models which may control for some unmeasured confounding along with that due to covariates.	Issues of residual confounding and reverse causation (endogeneity) may remain
3. Propensity score matching	Using selection on observables to estimate the probability of exposure or treatment conditioned on measured confounders	Potential issues of residual and unmeasured confounding
4. Difference-in-Differences, Regression discontinuity, Instrumental variables	Using exogenous variation occurring 'naturally' in the data to estimate causal effect	Rely on strong assumptions (e.g. time-invariant confounding, continuity of the assignment variable continuity, association of the instrument with the outcome exclusively through the treatment variable) which are difficult to test although various analyses may give additional support. Some potential issues of unmeasured confounding remain.
5. Randomised Controlled Trial/Random assignment	Exploit random assignment of individuals to a treatment and a control group to effectively account for sources of selection bias	Considered as the gold standard for the identification of causal effects

A large part of the evidence base on the impacts of benefit sanction (n = 74; 68%) reports labour market outcomes, such as welfare benefit exits and re-entry, employment status and transitions, duration of unemployment and earnings. Wider impacts appear in just 35 studies (32%) (Table 3).

Very few of the studies reporting wider impacts are of the experimental or quasi-experimental types (just 15 per cent in Types 4 and 5). More than half (51%) are Type 1 with most of the rest (34%) from Type 2. For studies focussing on labour market outcomes, one quarter (25%) are from experimental or quasi-experimental groups. Of the remainder, the largest group (34%) comes from Type 2 containing more complex regression-based approaches but a significant number are still more basic Type 1 studies (23%).

Table 3 Study design typology by labour market and wider outcomes

Study design typology	Labour market outcomes		Wider outcomes		Total (n)
	n	%	n	%	
1. Descriptive analysis, OLS, logistic/probit regression	19	23	18	51	37
2. Survival models, hierarchical models, time series, fixed effects	28	34	12	34	40
3. Propensity score matching	15	18	0	0	15
4. Difference-in-Differences, Regression discontinuity, Instrumental variables	9	11	3	9	12
5. Randomised Controlled Trial/Random assignment	11	14	2	6	13
Total	74	100	35	100	109

Note: Total in each column exceeds the number of studies reviewed reported here due to multiple outcomes in individual publications.

Given that studies of wider impacts have had little attention in the past, we focus for the remainder of the paper on the 35 studies in this group. We are interested in the evidence they contain on the wider impacts and also, insofar as the few studies we have can support this, the extent to which those with more sophisticated designs reinforce the findings of non-experimental studies or challenge them.

Table A3 in the Appendix provides an overview of the 35 studies reporting wider outcomes. All studies using quasi-experimental or experimental designs (Types 4 and 5) were from the US and all the studies included in Type 4 applied a difference-in-differences approach. Three studies focussed on child-related outcomes (Aber, Brooks-Gunn, and Maynard 1995, Fein and Lee 2003, Wang 2015), such as child development, well-being and maltreatment (e.g. foster care placement), and found no significant impacts of benefit sanctions. Two of these studies were based on random assignment of individuals to a control and a treatment group (Aber, Brooks-Gunn, and Maynard 1995, Fein and Lee 2003), while the third study by Wang (2015) combined a difference-in-differences model with a propensity score matching approach. For adult outcomes, only marginally significant increases were reported for school/training attendance by Aber and co-authors' (1995) experimental study and a significant reduction in the number of people claiming benefits was found in a quasi-experimental study conducted by Danielson and Klerman (2008), who applied a difference-in-differences estimation approach. No significant impacts were reported for demographic outcomes such as childbearing (Aber, Brooks-Gunn, and Maynard 1995) and living arrangements (Acs and Nelson 2004).

Among the studies included in Type 2 (enhanced regression models), a high proportion applied either fixed effect models, survival models, or a combination of the two. More than half of the studies focussed on child-related outcomes, such as maltreatment (Beimers and

Coulton 2011; Ovwigho, Leavitt, and Born 2003; Paxson and Waldfogel 2003; Slack, Lee, and Berger 2007), living arrangements (Dunifon, Hynes, and Peters 2009), and well-being (Lohman et al. 2004, Reichman, Teitler, and Curtis 2005). Benefit sanctions were found to have a positive association with some aspects of child maltreatment and a negative association with children's well-being, while no significant associations were found with children's living arrangements. Divergent associations were found for adult health outcomes, with a negative link reported by Davis (2019) and a positive link reported by Reichman and co-authors (2015). No effects of benefit sanctions were reported for demographic outcomes such as female headship (Fitzgerald and Ribar 2004) and non-marital childbearing (Ryan, Manlove, and Hofferth 2006). Significant associations were found by two ecological studies, with a study by Snarr (2013) showing that sanctions were linked to a reduction of welfare caseload (number of people claiming benefits), while Reeves and Loopstra (2017) reported that sanctions were correlated with a higher proportions in the population of people with a disability and lone parents. A study by Reichman and co-authors (2005) showed that benefit sanctions were positively linked with various aspects of material hardship, such as food deprivation, utility shutoffs and housing problems.

Among the studies included in Type 1, there is a balanced distribution among descriptive analyses and standard regression-based analyses using either OLS or logistic models, with equal numbers across the three sub-groups. Based on the US, the vast majority of studies (n = 13) reported heightened associations with various aspects of material hardship experienced by both adult and children, including financial strain, food deprivation, utility shutoffs, housing problems, difficulties in accessing medical care and perceived hardship. In an ecological study using OLS regression, Rodgers and co-authors (2006) found that harsher sanctions correlate with a reduction in poverty. No significant associations were reported in the case of adult's mental and physical health (Casey et al. 2004; Lindhorst and Mancoske 2006). Concerning welfare caseload, while a study by Chavkin and co-authors (2000) reported no significant associations with benefit sanctions, Yu (2001) revealed that sanctions programmes led to a significant caseload reduction. The remaining studies focussed on children and reported an increase in problematic outcomes concerning cognitive and behavioural aspects, hospitalisations, school enrolment and attendance (Chase-Lansdale et al 2002; Cook et al. 2002; Gritz et al 2001; Larson, Singh, and Lewis 2001).

4. Conclusion

In this study we presented preliminary findings from a scoping review that was conducted to identify the nature of the evidence base relating to impact of benefit sanctions on both labour market and wider outcomes. The review applied comprehensive searching of the international quantitative literature and rigorous methodology in line with the PRISMA guidelines designed for scoping reviews (Tricco et al. 2018). From the examination of the studies included in the scoping review, we developed a study design typology based on the extent to which the modelling approaches employed provide support for the identification of causal effects.

Our scoping review identified 109 studies providing novel quantitative evidence on the labour market and/or wider impacts of sanctions which met our inclusion criteria. In terms of geographical coverage, most of the studies originated from the US, followed by Western and Northern European studies. While the overall volume of studies has decreased in recent years, an increasing proportion are using experimental or quasi-experimental study designs which enables stronger claims to identify causal effects. Two thirds of our sample investigated labour market impacts while only one third focussed on the wider impacts. Of these only a small number used experimental or quasi-experimental designs. The nature of the impacts appeared rather fragmented, encompassing a wide range of outcomes relating to both adults and children. In both cases, experimental and quasi-experimental designs reported either non-significant or marginally significant effects. The results from non-experimental designs were highly inconsistent.

Future developments will include an in-depth analysis and synthesis of the evidence of benefit sanctions on labour market outcomes which will facilitate a comparison between the nature of the evidence base and study design features for both labour market and wider outcomes. In this study, analyses did not cover other components of the analytical sample, such as studies based on the assessment of multiple interventions or policy tools and studies including both narrative and more systematic reviews. It would be useful to extend the analyses to both groups of studies in order to identify whether there is any variation in terms of patterns emerging from the evidence base and related study design approaches. There is also scope to extend the explorative exercise undertaken for this scoping review to a full-systematic review, by conducting a critical appraisal of the evidence base by means of a more formal synthesis using a meta-analytic approach.

References

- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics. An empiricist's companion*. Princeton: Princeton University Press.
- Angrist, J. D. and Pischke, J.-S. (2015). *Mastering metrics. The path from cause to effect*. Princeton: Princeton University Press.
- Arksey, H., and O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8 (1): 19-32.
- Bonoli, G. (2010). The political economy of active labor-market policy. *Politics & Society*, 38(4): 435-457.
- Baumberg Geiger, B (2017). Benefits conditionality for disabled people: Stylized facts from a review of international evidence and practice. *Journal of Poverty and Social Justice*, 25(2): 107-128.
- Bonoli, G., and Natali, D. (2012). *The politics of the new welfare state*. Oxford: Oxford University Press.
- Card, D., Kluge, J, & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *Economic Journal*, 120: F452-F477.
- Card, D., Kluge, J, & Weber, A. (2018). What works? A meta analysis of recent active labour market policy evaluations. *Journal of the European Economic Association*, 16(3): 894-931.
- Craig, P., Katikireddi, S. V., Leyland, A., & Popham, F. (2017). Natural experiments: An overview of methods, approaches, and contributions to public health intervention research. *Annual Review of Public Health*, 38: 39-56.
- Dunning, T. (2012). *Natural experiments in the social sciences: A designed-based approach*. Cambridge: Cambridge University Press.
- Dwyer, P., and Wright, S. (2014). Universal Credit, ubiquitous conditionality and its implications for social citizenship. *Journal of Poverty and Social Justice*, 22(1): 27-39.
- Griggs, J., and Evans, M. (2010) *Sanctions within conditional benefit systems. A review of evidence*. York: Joseph Rowntree Foundation.
<https://www.jrf.org.uk/sites/default/files/jrf/migrated/files/conditional-benefit-systems-full.pdf>. Accessed 12 Jul 2019.
- Levac, D., Colquhoun, H., and O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation Science*, 5(69): 1-9.
- Murnane, R. J., and Willett, J. B. (2010). *Methods matter. Improving causal inference in educational and social science research*. Oxford: Oxford University Press.
- Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., and Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3): 141-46.
- Pattaro, S., Bailey, N, Williams, E, Gibson, M, Wells, V., Wright, S., and Tranmer, M. (2019) Labour market and wider impacts of benefit sanctions: A scoping review [Review Protocol]. Unpublished document. <http://eprints.gla.ac.uk/182309/7/182309.pdf>. Accessed 12 Jul 2019.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., Weeks, L., Hempel, S. et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, 169 (7): 467-473.

Vooren, M., Haelermans, C., Groot, W., & Maassen van den Brink, H. (2019). The effectiveness of active labor market policies: A meta-analysis. *Journal of Economic Surveys*, 33(1): 125-149.

References of studies included in the sample

Ala-Kauhala, M., and Parpo, A. (2012). Unemployment and the effects of activation policy on quality of life and self-performance. In T. Kieselbach, and Mannila, S. (eds.) *Unemployment, precarious work and health. Research and policy issues*. Springer. 379-392.

Aber, J. L., Brooks-Gunn, and Maynard, R. A. (1995). Effects of welfare reform on teenage parents and their children. *The Future of Children. Critical Issues for Children and Youths*, 5(2): 53-71.

Acs, G., and Nelson, S. (2004). Changes in living arrangements during the late 1990s: Do welfare policies matter? *Journal of Policy Analysis & Management*, 23(2): 273-290.

Beimers, D., and Coulton, J. (2011). Do employment and type of exit influence child maltreatment among families leaving Temporary Assistance for Needy Families? *Children and Youth Services Review*, 33(7): 1112-1119.

Casey, P., Goolsby, S., Berkowitz, C., Frank, D., Cook, J., Cutts, D., Black, M. M., Zaldivar, N., Levenson, S., Heeren, T., and Meyers, A. (2004). Maternal depression, changing public assistance, food security, and child health status. 113(2): 298-304.

Chase-Lansdale, P. L., Coley, R. L., Lohman, B. J., Pittman, L. D. (2002). *Welfare reform: What about the children? Welfare, children & families: A three-city study. Policy brief*. Baltimore: John Hopkins University. <https://files.eric.ed.gov/fulltext/ED464169.pdf>. Accessed 16 Jun 2019.

Chavkin, W., Romero, D., and Wise, P. H. (2000), State welfare reform policies and declines in health insurance. *American Journal of Public Health*, 90(6): 900-908.

Cherlin, A. J., Bogen, K., Quane, J. M., Burton, L. (2002). Operating within the rules: Welfare recipients' experiences with sanctions and case closings. *Social Service Review*, 76(3): 387-405.

Cook, J. T., Frank, D. A., Berkowitz, C., Black, M. M., Casey, P. H., Cutts, D. B., Meyers, A. F., Zaldivar, N., Skalicky, A., Levenson, S., and Heeren, T. (2002). Welfare reform and the health of young children: A sentinel survey in 6 US cities. *Archives of Pediatrics and Adolescent Medicine*, 156: 678-684.

Danielson, C., and Klerman, J. A. (2008). Did welfare reform cause the caseload decline? *Social Service Review*, 82(4): 703-730.

Davis, O. (2019). What is the relationship between benefit conditionality and mental health? Evidence from the United States on TANF policies. *Journal of Social Policy*, 48(2): 249-269.

Dunifon, R., Hynes, K., and Peters, E. H. (2009). State welfare policies and children's living arrangements. *Social Service Review*, 83(3): 351-388.

- Eardley, T. (2006). The Impact of breaching and financial penalties on income support recipients. *The Economic and Labour Relations Review*, 17(1): 79-105.
- Fein, D. J., and Lee, W. L. (2003). The impacts of welfare reform on child maltreatment in Delaware. *Children and Youth Services Review*, 25(1/2): 83-111.
- Fitzgerald, J. M., and Ribar, D. C. (2004). Welfare reform and female headship. *Demography*, 41(2): 189-212.
- Gritz, R. M., Mancuso, D. C., Lieberman, C. J., Lindler, V. L. (2001). Assessing the family circumstances of TANF applicants and leavers in Contra Costa and Alameda Counties. Final report. Burlingame, CA: The SPHERE Institute.
<https://aspe.hhs.gov/system/files/pdf/177086/AspeCCAFinal.pdf>. Accessed 14 April 2019.
- Hunter, T., and Santhiveeran, J. (2005). Experiences of material hardships among TANF leavers. *Journal of Family Social Work*, 9(1): 1-16.
- Kalil, A., Seefeldt, K. S., and Wang, H.-C. (2002). Sanctions and material hardship under TANF. *Social Service Review*, 76(4): 642-662.
- Lindhorst, T., Mancoske, R. J., and Kemp, A. A. (2000). Is Welfare Reform Working? A Study of the Effects of Sanctions on Families Receiving Temporary Assistance to Needy Families. *Journal of Sociology and Social Welfare*, 27(4): 171-184.
- Lindhorst, T., and Mancoske, R. J. (2006). The social and economic impact of sanctions and time limits on recipients of Temporary Assistance to Needy Families. *Journal of Sociology and Social Welfare*, 33(1): 93-114.
- Larson, A. M., Singh, S., and Lewis, C. (2011). Sanctions and education outcomes for children in TANF families. *Child & Youth Services*, 32(2): 180-189.
- Lee, B. J., Slack, K. S., and Lewis, D. A. (2004). Are welfare sanctions working as intended? Welfare receipts, work activity, and material hardship among TANF-recipient families. *Social Service Review*, 78(3): 370-403.
- Livermore, M., Powers, R. S., Lim, Y., Davis, B. C. (2015). Predicting material hardship among former welfare-to-work participants: An income- and resource-packaging model. *Journal of Social Service Research*, 41(2): 158-171.
- Lohman, B. J., Pittman, L. D., Coley, R. L., and Chase-Lansdale, P. L. (2004). Welfare history, sanctions, and developmental outcomes among low-income children and youth. *Social Service Review*, 78(1): 41-73.
- Oggins, J., and Fleming, A. (2001). Welfare reform sanctions and financial strain in a food-pantry sample. *Journal of Sociology and Social Welfare*, 28(2): 101-124.
- Ovwigbo, P. C., Leavitt, K. L., and Born, C. E. (2003) Risk factors of child abuse and neglect among former TANF families: Do later leavers experience greater risk? *Children and Youth Services Review*, 25(1/2): 139-163.
- Paxson, C., and Waldfogel, J. (2003). Welfare reforms, family resources, and child maltreatment. *Journal of Policy Analysis and Management*, 22(1): 85-113.
- Reeves, A., and Loopstra, R. (2017) 'Set up to fail'? How welfare conditionality undermines citizenship for vulnerable groups. *Social Policy and Society*, 16(2): 327-338.

- Reichman, N. E., Teitler, J. O., and Curtis, M. A. (2005). TANF sanctioning and hardship. *Social Service Review*, 79(2): 215-236.
- Rodgers, H. R., Payne, L. Jr., and Chervachidze, S. (2006). State Poverty Rates: Do the New Welfare Policies Make a Difference? *Review of Policy Research*, 33(3): 657-679.
- Ryan, S., Manlove, J., Hofferth, S. L. (2006). State-level welfare policies and nonmarital subsequent childbearing. *Population Research and Policy Review*, 25(1): 103-126.
- Slack, K. S., Lee, B. J., Berger, L. M. (2007). Do welfare sanctions increase child protection system involvement? A cautious answer. *Social Service Review*, 81(2): 207-228.
- Snarr, H. W. (2013). Was it the economy or reform that precipitated the steep decline in the US welfare caseload? *Applied Economics*, 45(4): 525-540.
- Wang, J. S.-H. (2015). TANF coverage, state TANF requirement stringencies, and child well-being. *Children and Youth Services Review*, 53: 121-129.
- Yu, H. (2001). Welfare reform and caseload reductions. *International Journal of Social Economics*, 28(4): 383-343.

Appendix

Table A1 Summary of search strategies and records retrieved

Database name and coverage	Interface	Database dates	Records retrieved
ASSIA	Proquest	1987-	324
BEI (British Education Index)	EBSCO	1929-	56
EconLit ^a	EBSCO	1886-	1033
EconLit ^b	EBSCO	1886-	1424
ERIC	EBSCO	1966-	492
PsycINFO	EBSCO	1698-	821
MEDLINE ^c	OVID	1996-	534
SCOPUS	Elsevier	2004-	2365
SocINDEX	EBSCO	1908-	2580
Total			9629
After duplicates removed			7169

Notes: ^a The searches for both this database and all the remaining unmarked databases, including Medline, were conducted in March 2019; ^b A revised search for EconLit was conducted in June 2019 in order to integrate the search terms relating to 'unemployment insurance' which were not previously included; ^c No revisions from 1996 to February Week 4 2019.

Table A2 Search strategies

Search	Terms
	ASSIA
Set 1	(noft(sanction*) OR noft(penalt*) OR noft(punishment*) OR noft(punitive) OR noft(monitoring) OR noft(exclusion) OR noft(leaving) OR noft(exiting)) OR noft(exit) OR noft(austerity) OR noft(conditionality) OR noft("welfare conditionality") OR noft("welfare sanction"))
Set 2	(noft(claimant*) OR noft("job seeker*") OR noft(unemployed) OR noft("welfare recipient*") OR noft(recipient*) OR noft(unemployment) OR noft(family) OR noft(families) OR noft(child*) OR noft(youth) OR noft(jobless) OR noft(sick) OR noft(sickness) OR noft(disabled) OR noft(disability) OR noft(impaired) OR noft(incapacity) OR noft(parent*) OR noft(lone) OR noft(single))
Set 3	(noft("TANF") OR noft("welfare to work") OR noft("public assistance") OR noft("employment and support allowance") OR noft("individual re-integration agreement") OR noft(monetary NEAR/2 incentive) OR noft("monetary N/3 incentive") OR noft("monetary benefit") OR noft("social assistance") OR noft("work first strateg*") OR noft("incapacity benefit*") OR noft("disability living allowance") OR noft("Personal Responsibility and Work Opportunity Reconciliation Act") OR noft("Active labor market polic*") OR noft("Active labor market program*") OR noft("agenda 2010") OR noft("domestic purposes benefit") OR noft("america works") OR noft("cash benefit*") OR noft("cash incentive") OR "government intervention*" OR noft("government program*") OR noft("income benefit*") OR noft("income supplement*") OR noft("job seeker*") OR noft("job seekers allowance") OR noft("public welfare reform*") OR noft("tax credit*") OR noft("universal credit") OR noft("financial benefit*") OR noft("conditional benefit*") OR noft("work program*") OR noft("social security reform*") OR noft("claimant*") OR noft("welfare reform*") OR noft("benefit cap") OR noft("welfare conditionality") OR noft("social protection scheme*") OR noft("temporary assistance to needy families"))
Set 4	All sets combined
	EBSCO – for all databases
S1	"help to work program"
S2	"canada health and social transfer"
S3	"Personal Responsibility and Work Opportunity Reconciliation Act"
S4	"Active labor market polic*"
S5	Active labor market program* OR active labor participation
S6	"Agenda 2010"
S7	"domestic purposes benefit"
S8	"America works"
S9	cash benefit*
S10	cash incentives
S11	"government intervention*"
S12	"government program*"
S13	"income benefit"
S14	"income support"
S15	"income supplement*"
S16	"job seeker allowance*"
S17	"public welfare reform*"
S18	"tax credit"

Table A2 (Continued)

Search	Terms
S19	"universal credit"
S20	"financial benefit*"
S21	"conditional benefits"
S22	"work program*"
S23	"social security reform*"
S24	"welfare reform*"
S25	"benefit cap"
S26	welfare conditionality
S27	((DE "PUBLIC welfare policy") OR (DE "WELFARE recipients")) OR (DE "WELFARE state")
S28	"public assistance"
S29	DE "SOCIAL security"
S30	"social protection scheme"
S31	"Temporary Assistance to Needy Families" OR tanf
S32	"welfare to work"
S33	"Employment and Support Allowance"
S34	hilfe zum arbeit OR hilfe zum lebensunterhalt
S35	"individual re-integration agreement"
S36	monetary N3 incentive
S37	"monetary benefit*"
S38	"social assistance"
S39	"work first strateg*"
S40	"disability living allowance"
S41	basic income
S42	welfare funds
S43	ontario works
S44	mandatory employment
S45	new start allowance
S46	jobbskatteavdraget
S47	workfare
S48	disability benefit*
S49	incapacity benefit*
S50	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49
S51	sanction*
S52	penalt*
S53	punishment*
S54	punitive
S55	welfare conditionality
S56	"welfare sanction"
S57	austerity
S58	exclusion OR exit* OR leaving or loss
S59	monitoring
S60	S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59

Table A2 (Continued)

Search	Terms
S61	claimant* or parent* or family or families or child* or youth or lone or single or disabled or disability or impaired or incapacity or sick or sickness
S62	job seeker*
S63	jobless*
S64	recipient*
S65	"welfare recipient"
S66	unemployed or unemployment
S67	S61 OR S62 OR S63 OR S64 OR S65 OR S66
Medline	
1	"welfare benefits".ab,ti.
2	help to work program.ab,ti.
3	(canada health and social transfer).ab,ti.
4	(Personal Responsibility and Work Opportunity Reconciliation Act).ab,ti.
5	"Active labor market polic*".ab,ti.
6	labor force participation.ab,ti.
7	"Active labor market program*".ab,ti.
8	Agenda 2010.ab,ti.
9	domestic purposes benefit.ab,ti.
10	america works.ab,ti.
11	cash benefit.ab,ti.
12	cash incentive.ab,ti.
13	government intervention.ab,ti.
14	"Government program*".ab,ti.
15	"income benefit*".ab,ti.
16	income support.ab,ti.
17	income supplement.ab,ti.
18	job seeker allowance.ab,ti.
19	"public welfare reform*".ab,ti.
20	"tax credit*".ab,ti.
21	basic income.ab,ti.
22	universal credit.ab,ti.
23	financial benefit*.ab,ti.
24	conditional benefits.ab,ti.
25	"work program*".ab,ti.
26	social security reform*.ab,ti.
27	"welfare reform*".ab,ti.
28	welfare fund.ab,ti.
29	benefit cap.ab,ti.
30	welfare conditionality.ab,ti.
31	Social Welfare/
32	public assistance/
33	Social Security/
34	"social protection scheme*".ab,ti.
35	Temporary Assistance to Needy Families.ab,ti.
36	TANF.ab,ti.

Table A2 (Continued)

Search	Terms
37	welfare to work.ab,ti.
38	public assistance.ti,ab.
39	(Employment and Support Allowance).ab,ti.
40	hilfe zum arbeit.ab,ti.
41	hilfe zum lebensunterhalt.ab,ti.
42	revenu minimum d'insertion.ab,ti.
43	individual re-integration agreement.ab,ti.
44	(monetary adj 2 incentive*).ab,ti.
45	"monetary benefit*".ab,ti.
46	monetary support.ab,ti.
47	social assistance.ab,ti.
48	"work first strateg*".ab,ti.
49	"incapacity benefit*".ab,ti.
50	disability living allowance.ab,ti.
51	disability benefit*.ab,ti.
52	workfare.ab,ti.
53	new start allowance.ab,ti.
54	jobbskatteavdraget.ab,ti.
55	mandatory employment.ab,ti.
56	Ontario works.ab,ti.
57	"sanction*".ab,ti.
58	"penalt*".ab,ti.
59	"punishment*".ab,ti.
60	punitive.ab,ti.
61	welfare conditionality.ab,ti.
62	conditionality.ab,ti.
63	"welfare sanction*".ab,ti.
64	austerity.ab,ti.
65	loss.ab,ti.
66	exclusion.ab,ti.
67	exit*.ab,ti.
68	monitoring.ab,ti.
69	leaving.ab,ti.
70	"job seeker*".ab,ti.
71	jobless.ab,ti.
72	"recipient*".ab,ti.
73	"welfare recipient*".ab,ti.
74	unemployed.ab,ti.
75	unemployment.ab,ti.
76	(claimant* or parent* or family or families or child* or youth or lone or single or disabled or disability or impaired or incapacity or sick or sickness).ab,ti.
77	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56

Table A2 (Continued)

Search	Terms
78	57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69
79	70 or 71 or 72 or 73 or 74 or 75 or 76
80	77 and 78 and 79
SCOPUS	
1	(((TITLE-ABS-KEY ("Ontario Works") OR TITLE-ABS-KEY ("welfare benefits") OR TITLE-ABS-KEY ("exit to work") OR TITLE-ABS-KEY ("unemployment insurance") OR TITLE-ABS-KEY ("Help to work program*") OR ("Personal Responsibility and Work Opportunity Reconciliation Act") OR (TITLE-ABS-KEY ("Active labor market polic*") OR (TITLE-ABS-KEY (active AND labor AND market AND program*) OR (TITLE-ABS-KEY ("Agenda 2010") OR (TITLE-ABS-KEY ("the new deal") OR (TITLE-ABS-KEY (labor AND force AND participation) OR (TITLE-ABS-KEY ("domestic purposes benefit"))) OR ((TITLE-ABS-KEY ("America works") OR (TITLE-ABS-KEY (cash AND benefit*) OR (TITLE-ABS-KEY (cash AND incentive*) OR (TITLE-ABS-KEY ("government intervention*") OR (TITLE-ABS-KEY ("government program*") OR (TITLE-ABS-KEY ("income benefit") OR (TITLE-ABS-KEY ("income support") OR (TITLE-ABS-KEY ("income supplement*") OR (TITLE-ABS-KEY ("job seeker allowance*"))) OR ((TITLE-ABS-KEY ("tax credit") OR (TITLE-ABS-KEY ("universal credit") OR (TITLE-ABS-KEY ("financial benefit*") OR (TITLE-ABS-KEY ("conditional benefits") OR (TITLE-ABS-KEY ("work program*") OR (TITLE-ABS-KEY (workfare) OR (TITLE-ABS-KEY ("social security reform*") OR (TITLE-ABS-KEY ("welfare reform*") OR (TITLE-ABS-KEY ("benefit cap"))) OR ((TITLE-ABS-KEY (welfare AND conditionality)) OR (TITLE-ABS-KEY ("public assistance") OR (TITLE-ABS-KEY ("social protection scheme") OR (TITLE-ABS-KEY ("Temporary Assistance to Needy Families" OR tanf) OR (TITLE-ABS-KEY ("welfare to work") OR (TITLE-ABS-KEY ("Employment and Support Allowance") OR (TITLE-ABS-KEY (hilfe AND zum AND arbeit OR hilfe AND zum AND lebensunterhalt) OR (TITLE-ABS-KEY ("individual re-integration agreement") OR (TITLE-ABS-KEY ("monetary benefit*") OR (TITLE-ABS-KEY ("monetary incentive*"))) OR ((TITLE-ABS-KEY ("social assistance") OR (TITLE-ABS-KEY ("work first strateg*") OR (TITLE-ABS-KEY ("incapacity benefit") OR (TITLE-ABS-KEY ("disability living allowance") OR (TITLE-ABS-KEY ("Canada health and social transfer") OR (TITLE-ABS-KEY ("Active labor force participation)) or (TITLE-ABS-KEY(public welfare reform*)) or (TITLE-ABS-KEY(basic income)) or (TITLE-ABS-KEY(welfare fund*)) (TITLE-ABS-KEY(" revenue AND minimum AND d'insertion ")) or (TITLE-ABS-KEY(" monetary AND support ") or (TITLE-ABS-KEY(" disability AND living AND allowance*) OR (TITLE-ABS-KEY ("disability benefit") OR (TITLE-ABS-KEY ("new start allowance") OR (TITLE-ABS-KEY ("jobbskatteavdraget") OR (TITLE-ABS-KEY ("mandatory employment"))))
2	AND (TITLE-ABS-KEY (impaired OR incapacity OR parent* OR lone OR single OR sickness OR sick OR family OR families OR child* OR youth OR claimant* OR "job seeker" OR unemployment OR unemployed OR "welfare recipient*" OR jobless* OR recipient* OR disability OR disabled)
3	AND (TITLE-ABS-KEY (sanction* OR punishment* OR penalt* OR punitive OR exit* OR monitoring OR leaving OR exclusion OR austerity OR loss OR conditionality OR "welfare conditionality" OR "welfare sanction")

Table A2 (Continued)

Search	Terms
	EconLit (revised search conducted in June 2019)
S1	"help to work program"
S2	"canada health and social transfer"
S3	"Personal Responsibility and Work Opportunity Reconciliation Act"
S4	"Active labor market polic*"
S5	"Agenda 2010"
S6	"domestic purposes benefit"
S7	"America works"
S8	"cash benefit*"
S9	"cash incentive*"
S10	"government intervention*"
S11	"government program*"
S12	"income benefit*"
S13	"income support"
S14	"income supplement*"
S15	"job seeker allowance*"
S16	"public welfare reform*"
S17	"tax credit*"
S18	"universal credit"
S19	"financial benefit*"
S20	"conditional benefits"
S21	"work program*"
S22	"welfare reform*"
S23	"benefit cap"
S24	"welfare conditionality"
S25	"Temporary Assistance to Needy Families" OR "TANF"
S26	"welfare to work"
S27	"Employment and Support Allowance"
S28	hilfe zum arbeit OR hilfe zum lebensunterhalt
S29	"individual re-integration agreement"
S30	"monetary benefit*"
S31	"social assistance"
S32	"work first strateg*"
S33	"disability living allowance"
S34	"basic income"
S35	"basic income guarantee"
S36	"ontario works"
S37	"mandatory employment"
S38	"new start allowance"
S39	jobbskatteavdraget
S40	workfare
S41	"disability benefit*"
S42	"incapacity benefit*"

Table A2 (Continued)

Search	Terms
S43	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42
S44	(ZU "unemployment insurance") or (ZU "unemployment insurance; severance pay; plant closings") or (ZU "unemployment assistance")
S45	(ZU "welfare, well-being, and poverty: government programs; provision and effects of welfare programs" OR ZU "social security"))
S46	S44 OR S45
S47	S43 OR S46
S48	AB sanctions* OR TI sanction*
S49	AB penalt* OR TI penalt
S50	AB punishment* OR TI punishment*
S51	AB punitive OR TI punitive
S52	AB conditionality OR TI conditionality
S53	AB austerity OR TI austerity
S54	TI exclusion OR AB exclusion
S55	AB exit* AND TI exit*
S56	AB leaving OR TI leaving
S57	AB loss AND TI loss
S58	TI monitoring OR AB monitoring
S59	S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58
S60	S47 OR S59

Table A3 Overview of studies on the wider impacts of benefit sanctions included in the sample

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
Study design typology - Type 5: Randomised controlled trial, randomised assignment (Experimental design)									
1	Aber, Brooks-Gunn, and Maynard (1995) USA (three cities, Illinois and New Jersey)	Lone parents (teenage parents with one child)	<p><i>Outcomes:</i> <u>Adult outcomes:</u> School attendance Childbearing <u>Parent/child outcomes:</u> Parenting behaviour Child development</p> <p><i>Exposure:</i> Loss of mother's component of welfare benefit</p>	Experimental (random assignment)	<p><i>Findings:</i> <u>Adult:</u> Attendance at school/job training increased. <u>Parent/child:</u> No reduction in further childbearing or change in parenting behaviour or child development. "The evaluation results suggest that supportive, mandatory welfare-to-work interventions need not harm parents or their children in the short term, and that their modest positive effects on the financial independence of the teenage mothers may yield long-term rewards" (p. 53).</p> <p><i>Time horizon:</i> Short term.</p>	Linked survey-administrative data (various sources)	<p><u>Adults:</u> 4559 (admin data outcomes); 3867 (survey outcomes). <u>Parenting/child:</u> 182 (mother-child pairs at one site)</p>	n/a	n/a
2	Fein and Lee (2003) USA (Delaware)	Lone parents	<p><i>Outcomes:</i> <u>Child outcomes:</u> child maltreatment (neglect, physical and emotional abuse)</p> <p><i>Exposure:</i> Partial to full benefit sanctions</p>	Experimental (random assignment)	<p><i>Findings:</i> "Results show small increases in child neglect but no effects on physical abuse, sexual abuse, or foster care placement" (p. 83).</p> <p><i>Time horizon:</i> Medium term.</p>	Linked survey-administrative data (various sources)	3959	Control group had significant lower rates of work participation, experience of any sanctions and full sanctions.	n/a

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
Study design typology - Type 4: Difference-in-Differences, Regression discontinuity, Instrumental variables (Quasi-experimental design)									
3	Acs and Nelson (2004) USA (13 states)	Low-income families	<i>Outcomes:</i> <u>Demographic outcomes:</u> Dual versus single parenting) <i>Exposure:</i> Full-family sanctions	Quasi-experimental (Difference-in-difference-in-differences)	<i>Findings:</i> “Sanctions [...] have no clear consistent association with living arrangements” (p. 273). <i>Time horizon:</i> Short/medium term.	National Surveys of America’s Families (NSAF) 1997, 1999	n/a	Same states but less poor or slightly better educated	Yes
4	Danielson and Klerman (2008) USA	Low-income families	<i>Outcomes:</i> Welfare caseload (all benefit recipients) <i>Exposure:</i> Gradual full-family sanctions	Quasi-experimental (Difference-in-differences with Fixed Effects)	<i>Findings:</i> The state-level introduction of full-family sanction policies reduces the welfare caseload three years after implementation. <i>Time horizon:</i> Long term.	Administrative data Welfare Rules Database 1990-2005	9359	n/a	Yes
5	Wang (2015) USA	Low-income families	<i>Outcomes:</i> <u>Child outcomes:</u> Child well-being (educational outcomes) <i>Exposure:</i> Gradual full-family sanctions	Quasi-experimental (Difference-in-differences; Propensity Score Matching)	<i>Findings:</i> “Neither state TANF policy stringencies nor income and employment changes show significant differential influence on TANF’s impact on child well-being” (p. 121). <i>Time horizon:</i> Medium term.	Survey of Income and Program Participation (SIPP), 2004 and 2008 panels	4163	Children not covered by TANF welfare policy tended to be from a white ethnic group, with a younger/less-educated guardian/parent	Yes

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
Study design typology – Type 2: Survival models, hierarchical models, time series, fixed effects (Non-experimental 2)									
6	Beimers and Coulton (2011) USA (Ohio)	Low-income families	<i>Outcomes:</i> <u>Child outcome:</u> Child maltreatment <i>Exposure:</i> Involuntary exit from welfare	Non-experimental (Survival model)	<i>Findings:</i> “Families with an involuntary exit from TANF were at increased risk of a substantiated or indicated finding of maltreatment” (p. 1112). <i>Time horizon:</i> Medium term.	Linked administrative data	18000	n/a	n/a
7	Davis (2019) USA	Lone parents (low-educated single mothers)	<i>Outcomes:</i> Mental health <i>Exposure:</i> Partial to full benefit sanctions (at state level)	Non-experimental (Fixed effects and Random effects models)	<i>Findings:</i> States that have harsher sanctions have worse mental health among low-educated single mothers. <i>Time horizon:</i> Short term.	Linked administrative data	233716	Mothers (excluded single mothers) who did not graduate from high school	Yes
8	Dunifon, Hynes, and Peters (2009) USA	Low-income families/lone parents	<i>Outcomes:</i> Child outcomes (living arrangements) <i>Exposure:</i> Partial to full benefit sanctions (at state level)	Non-experimental (Multinomial logistic regression with fixed effects)	<i>Findings:</i> State-level sanctions policies are not shown to affect children’s living arrangements. <i>Time horizon:</i> Short/medium term.	Linked data: Survey of Income and Programme Participation (SIPP) 1992, 1993, and 1996; Welfare Rules Database	45847	n/a	Yes

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
9	Fitzgerald, J. M. and Ribar, D. C. (2004) USA	Low-income families/lone parents	<i>Outcomes:</i> Demographic outcomes (Female headship) <i>Exposure:</i> Partial to full benefit sanctions (at state level)	Non-experimental (Survival model)	<i>Findings:</i> State-level sanctions policies are not shown to affect female headship. <i>Time horizon:</i> Medium term.	Linked data: Survey of Income and Programme Participation (SIPP) 1990, 1992, 1993, and 1996; Welfare Rules Database	By female headship-related outcomes: 654327 (level) 13822 (exits) 52839 (entry)	n/a	Yes
10	Lohman et al. (2004) USA (Boston, Chicago, San Antonio)	Low-income families	<i>Outcomes:</i> <u>Child outcomes:</u> Quantitative/reading skills, behavioural problems <i>Exposure:</i> Partial or full benefit sanction	Non-experimental (Hierarchical model)	<i>Findings:</i> “Results show that while sanctions are not linked to adolescents’ outcomes, there is a negative link between sanctions and young children’s well-being [in terms of lower cognitive achievement and serious behavioural problems]” (p. 67). <i>Time horizon:</i> Short/medium term.	Survey data: Welfare, Children, and Families: Three-City Study, 1999	1885	n/a	No
11	Ovwigo, Leavitt, and Born (2003) USA (Maryland)	Low-income families	<i>Outcomes:</i> Child maltreatment <i>Exposure:</i> Involuntary exit from welfare due to sanctions	Non-experimental (Survival model)	<i>Findings:</i> Those exiting welfare due to sanctions have significantly higher maltreatment rates than almost all other groups <i>Time horizon:</i> Medium term.	Linked administrative data	17440 children in 8900 families	n/a	n/a

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
12	Paxson and Waldfogel (2003) USA	Low-income families	<i>Outcomes:</i> <u>Child outcome:</u> Child maltreatment (Reports of child maltreatments, Cases of abuse/neglect, Children in foster care <i>Exposure:</i> Full-family sanctions (at state-level)	Non-experimental (Fixed effects models)	<i>Findings:</i> Sanctions are positively and significantly related to the number of children placed in foster care, but are not related to the other maltreatment measures. <i>Time horizon:</i> Short/medium term.	State-level administrative and survey data: (various sources) 1990-1998	49	n/a	Yes
13	Reeves and Loopstra (2017) UK (Great Britain)	Unemployed	<i>Outcomes:</i> Persons with disability, Lone parents <i>Exposure:</i> Proportion of benefit claimants who received a sanction (local-authority level)	Non-experimental (Fixed effects models)	<i>Findings:</i> “Sanction rates appear to be higher in areas where there are more disabled and lone parent JSA claimants. [...] Conditionality appears to disadvantage those with ill health, physical limitations, or uncertain family commitments” (p. 335). <i>Time horizon:</i> n/a.	Area-level administrative and survey data (various sources) 2008/09-2014/15	175	n/a	Yes

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
14	Reichman, Teitler, and Curtis (2005) USA	Lone parents	<p><i>Outcomes:</i> <u>Material hardship</u> (maternal/child hunger, homelessness/eviction, utility shutoffs, lack of medical care, any of above hardships, received financial support from family/ friends, moved in with family/ friends) <u>Mother's physical and mental health:</u> (Depression/anxiety, Self-reported physical health) <u>Child's physical health:</u> (reported by mother)</p> <p><i>Exposure:</i> Partial to full benefit sanctions</p>	Non-experimental (Logistic regression models with fixed effects)	<p><i>Findings:</i> Benefit sanctions have a significant positive association with hunger, utility shutoffs, material hardship, poor maternal physical health, and relying on others for housing.</p> <p><i>Time horizon:</i> Medium term.</p>	Fragile Families and Child Wellbeing Survey, 1998-2000	821	Non-sanctioned mothers	Yes
15	Ryan, Manlove, and Hofferth (2006) USA	Lone parents	<p><i>Outcomes:</i> <u>Demographic outcome:</u> Nonmarital childbearing</p> <p><i>Exposure:</i> Partial to full benefit sanctions (at state-level)</p>	Non-experimental (Survival model with fixed effects)	<p><i>Findings:</i> "[...] Sanction policies [...] do not have any influence on women's childbearing behaviors [...], net of women's individual characteristics and state economic environments." (p. 103).</p> <p><i>Time horizon:</i> Medium term.</p>	Panel Study of Income Dynamics (PSID), 1989-1996	458	Lone mothers who had no subsequent nonmarital birth tended to live in states with lower unemployment rates and higher median incomes compared to their counterparts	Yes

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
16	Slack, Lee, and Berger (2007) USA (Illinois)	Lone parents	<i>Outcomes:</i> <u>Child outcomes:</u> Child maltreatments (Reports of child neglect and abuse; Indicated reports of maltreatments) <i>Exposure:</i> Partial to full benefit sanctions (at state-level)	Non-experimental (Survival model with fixed effects)	<i>Findings:</i> "Receipt of [...] sanctions increases the rate of having an investigation for neglect but does not bear a statistically significant relation to having an indicated report of neglect or abuse" (p.207). <i>Time horizon:</i> n/a.	Linked data: Illinois Family Study (IFS) 1999-2000 Admin data (various sources)	1260	n/a	n/a
17	Snarr (2013) USA	Low-income families	<i>Outcomes:</i> <u>Material hardship:</u> Welfare caseload <i>Exposure:</i> Full-family sanctions (at state-level)	Non-experimental (Three-Stage Least Squares)	<i>Findings:</i> "Results indicate that welfare reform in general is the primary driver of the marked reduction in state caseloads between its peak in 1994 to 2005. The largest total effect is associated with full sanctions" (p.207). <i>Time horizon:</i> n/a.	Various sources, 1990-2005	n/a	n/a	Yes
Study design typology - Type 1: Descriptive analysis, OLS, logistic/probit regression (Non-experimental design 1)									
18	Ala-Kauhaluoma and Parpo (2012) Finland	Long-term unemployed, young unemployed	<i>Outcomes:</i> Self-performance and quality of life <i>Exposure:</i> Activation measures	Non-experimental (OLS regression)	<i>Findings:</i> "The activation plan [...] seems to have no clear impact on self-performance" (p. 387) <i>Time horizon:</i> Medium term.	Survey data	~ 2000	Unemployed people on waiting lists for activation measures	n/a
19	Casey et al. (2004) USA	Lone parents	<i>Outcomes:</i> Mental health (self-reported) <i>Exposure:</i> Partial to full benefit sanctions	Non-experimental (Logistic regression)	<i>Findings:</i> Benefit sanctions have no significant association with maternal depression. <i>Time horizon:</i> Medium term.	Children's Sentinel Nutritional Assessment Program, 2000-2001	5306	Benefit claimants whose benefits were not decreased	No

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/ control group (characteristics)	Inclusion of contextual indicators
Study design typology - Type 1: Descriptive analysis, OLS, logistic/probit regression (Non-experimental design 1)									
20	Chase-Lansdale et al. (2002) USA (Boston, Chicago, and San Antonio)	Low-income families	<i>Outcomes:</i> <u>Child well-being:</u> cognitive achievement, emotional/behavioural problems <i>Exposure:</i> Partial to full benefit sanctions	Non-experimental (Descriptive-survey frequencies)	<i>Findings:</i> Pre-schoolers and adolescents in sanctioned families showed problematic cognitive and behavioural outcomes. <i>Time horizon:</i> Medium term.	Survey data: Children and Welfare: a three-city study, 1999	1885	Claimants/recent leavers who have not been sanctioned	No
21	Chavkin, Romero, and Wise (2000) USA	Low-income families	<i>Outcomes:</i> Welfare caseload <i>Exposure:</i> Partial to full benefit sanctions (at state level)	Non-experimental (OLS regression)	<i>Findings:</i> State-level sanctions policies are not shown to be associated with welfare caseload <i>Time horizon:</i> Medium term.	Linked admin-survey data: various sources, 1995-1998	50	n/a	Yes
22	Cherlin et al. (2002) USA (Boston, Chicago, and San Antonio)	Low-income families	<i>Outcomes:</i> <u>Material hardship</u> (various indicators) <i>Exposure:</i> Partial to full benefit sanctions	Non-experimental (OLS regression)	<i>Findings:</i> “[...] Families that lose benefits tend to cut spending and rely mainly on friends and kin for support” (pp. 400-401) <i>Time horizon:</i> Medium term.	Survey data: Children and Welfare: a three-city study, 1999	1262	n/a	No
23	Cook et al. (2002) USA (6 cities)	Low-income families	<i>Outcomes:</i> <u>Child outcomes:</u> food security, health (lifetime hospitalisations, emergency department admissions) <i>Exposure:</i> Partial to full benefit sanctions	Non-experimental (Logistic regression)	<i>Findings:</i> Benefit sanctions are associated with a significant greater increase in the likelihood that young children will experience food insecurity and hospitalisations. <i>Time horizon:</i> Medium term.	Survey data: Children’s Sentinel Nutritional Assessment Program, 1998-2000	2718	Comparable households who had not been sanctioned	No

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
24	Eardley (2006) Australia	Unemployed	<i>Outcomes:</i> Compliance with requirements; <u>Material hardship:</u> Living expenses problems, housing problems, Borrowing/debt; Health; Social relationships Problems; Risk-taking behaviour <i>Exposure:</i> benefit sanctions	Non-experimental (OLS and logistic regression)	<i>Findings:</i> Sanctions lead to greater compliance with requirements but negative effects are nevertheless widespread, including impacts on living expenses, reliance on friends/family for support, housing problems, health impacts and risk-taking activities. <i>Time horizon:</i> Short term.	Survey data: Telephone survey, 2002	1005	n/a	n/a
25	Gritz et al. (2001) USA (California)	Low-income families	<i>Outcomes:</i> <u>Adult outcomes:</u> Health insurance, housing conditions, stable child care, substance abuse <u>Child outcome:</u> Child risk behaviours, <i>Exposure:</i> Welfare leavers	Non-experimental (descriptive analysis)	<i>Findings:</i> “The transition to child-only group [mainly leavers through sanctions] shows much less improvement than leavers and the informally diverted [i.e. other leavers] (p. xi)”. Worse on wide range of outcomes. <i>Time horizon:</i> Short/medium term.	Follow-up surveys of welfare leavers	~ 650	Voluntary welfare leavers	n/a

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
26	Hunter and Santhiveeran (2005) USA	Low-income families	<p><i>Outcomes:</i> <u>Material hardship:</u> Food deprivation, financial hardship, inability to pay rent/utility bills</p> <p><i>Exposure:</i> Partial to full benefit sanctions</p>	Non-experimental (Descriptive analysis)	<p><i>Findings:</i> "Voluntary leavers tended to experience housing related hardships more than involuntary leavers did. The majority of TANF leavers [...] reported worrying about food, which was ranked as the number one hardship. When compared to White leavers, other ethnic groups experienced a higher percentage of housing and healthcare-related hardships" (p. 1).</p> <p><i>Time horizon:</i> n/a.</p>	Survey data: National Survey of America's Families (NSAF), 1999	220	n/a	n/a
27	Kalil, Seefeldt, and Wang (2002) USA (Michigan)	Lone parents	<p><i>Outcomes:</i> <u>Material hardship:</u> Utility cut-off, had at least one hardship experience, expected hardship in the next 2 months</p> <p><i>Exposure:</i> Partial or full benefit sanction</p>	Non-experimental (Logistic regression)	<p><i>Findings:</i> Sanctions are associated with utility shutoffs, engaging in hardship-related activities, and subjective perceptions of material hardships</p> <p><i>Time horizon:</i> medium term</p>	Survey data: Women's Employment Survey, 1997, 1998 and 1999	562	Non-sanctioned single mothers were more likely to be cohabiting, non African American, younger and more highly educated	No
28	Larson, Singh, and Lewis (2011) USA (Minnesota)	Low-income families	<p><i>Outcomes:</i> <u>Child outcomes:</u> Education, enrolment disruptions</p> <p><i>Exposure:</i> Partial to full benefit sanctions</p>	Non-experimental (Descriptive analysis)	<p><i>Findings:</i> "Children from families with at least one sanction had lower mean attendance and significantly more disruptions to their school enrollment" (p. 194).</p> <p><i>Time horizon:</i> Short/medium term</p>	Linked administrative data (various sources)	19381	School-aged children in non-sanctioned low-income families were more likely to be white and in elementary school	n/a

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
29	Lee, Slack, and Lewis (2004) USA (Illinois)	Lone parents	<i>Outcomes:</i> <u>Material hardship:</u> Rent, utility, food, perceived overall hardship <i>Exposure:</i> Partial to full benefit sanctions	Non-experimental (Logistic regression)	<i>Findings:</i> “The findings show that TANF grant loss in general is associated with increased levels of hardship. Sanctions have a particularly strong relation to increased food hardships, while other types of grant loss are associated to a statistically significant degree with increased rent and utility hardships, and with perceived hardship” (p. 394). <i>Time horizon:</i> Short term.	Linked data: Illinois Family Study (IFS), 1999/2000, 2001, 2002; Admin data from Illinois Dept. of Employment Security	921	n/a	No
30	Lindhorst, Mancoske, and Kemp (2000) USA (Southern metropolitan region)	Low-income families	<i>Outcomes:</i> <u>Material hardship:</u> Various conditions, total number of problems <i>Exposure:</i> Full-family sanctions	Non-experimental (descriptive analysis)	<i>Findings:</i> Sanctioned families report an increase in the overall number of problems they are experiencing [...]. Compared to those who left voluntarily, those who were sanctioned off welfare were significantly different in terms of having unmet medical needs, going without food, and having their utilities turned off (p. 198)”. <i>Time horizon:</i> Medium term.	Survey data: Survey conducted by authors, 1998	347	Not sanctioned welfare leavers tend to be younger and have less number of problems	No
31	Lindhorst and Mancoske (2006) USA (Louisiana)	Lone parents	<i>Outcomes:</i> <u>Material hardship:</u> Food deprivation, housing problems, lacking medical care, no Medicaid receipt; <u>Physical and mental health:</u> Depression, Poor physical health <i>Exposure:</i> Sanctioned welfare leavers	Non-experimental (descriptive analysis)	<i>Findings:</i> “Sanction leavers [...] reported the highest levels of food and housing problems [and had difficulties in accessing medical resources], compared to TANF recipients (p. 9)”. <i>Time horizon:</i> Short/medium term.	Panel Study of Welfare Recipients, 1998-2001	277	n/a	No

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).

Table A3 Continued

No.	Author (year) Country	Target population	Wider outcomes and exposure	Study design (method)	Key findings and time horizon of effects ^a	Data source	Sample size	Comparator/control group (characteristics)	Inclusion of contextual indicators
32	Livermore et al. (2015) USA (Louisiana)	Lone parents	<i>Outcomes:</i> <u>Material hardship:</u> Food deprivation; Utility shut off; Housing problems (eviction/homelessness) <i>Exposure:</i> Sanctioned welfare leavers	Non-experimental (Logistic regression)	<i>Findings:</i> Those experiencing a sanction were more likely to experience material hardship. <i>Time horizon:</i> Medium term.	Survey data collected by authors, 2007; Admin data (various sources)	459	n/a	No
33	Oggins and Fleming (2001) USA (New York)	Low-income families	<i>Outcomes:</i> <u>Material hardship:</u> Various indicators <i>Exposure:</i> Benefit sanctions	Non-experimental (Logistic regression)	<i>Findings:</i> After being sanctioned, respondents noted greater hardship paying for adult health care, bills and rent; a greater likelihood of moving due to lack of rent; and less likelihood of having a phone-conditions that could make it even harder to work" (p. 116). <i>Time horizon:</i> n/a.	Survey conducted by authors, 1997 and 1999	118	n/a	No
34	Rodgers, Payne, and Chervachidze (2006) USA	Low-income families	<i>Outcomes:</i> <u>Material hardship:</u> Poverty rate <i>Exposure:</i> Weak, moderate, strong sanctions (at state-level)	Non-experimental (OLS regression)	<i>Findings:</i> "The evidence seems to suggest that tougher sanctions do tend to be effective in reducing poverty" (p. 674). <i>Time horizon:</i> n/a.	Various sources, 2003	n/a	n/a	Yes
35	Yu (2001) USA	Low-income families	<i>Outcomes:</i> <u>Material hardship:</u> Welfare caseload <i>Exposure:</i> Percentage of welfare recipients experiencing benefit reductions (at state-level)	Non-experimental (OLS regression)	<i>Findings:</i> There is a significant negative relationship between sanction programmes and welfare caseload. <i>Time horizon:</i> n/a.	Various sources	n/a	n/a	Yes

Note: ^a Time horizon of effects is classified as: Short-term (<6 months); Medium term (6-35 months), Long term (36+ months).