REVIEW ARTICLE

Check for updates

The effect of leave policies on increasing fertility: a systematic review

Jac Thomas [1✉], Francisco Rowe[1], Paul Williamson[1] & Eric S. Lin[2]

Low fertility is set to worsen economic problems in many developed countries, and maternity, paternity, and parental leave have emerged as key pro-natal policies. Gender inequity in the balance of domestic and formal work has been identified as a key driver of low fertility, and leave can potentially equalise this balance and thereby promote fertility. However, the literature contends that evidence for the effect of leave on fertility is mixed. We conduct the first systematic review on this topic. By applying a rigorous search protocol, we identify and review empirical studies that quantify the impact of leave policies on fertility. We focus on experimental or quasi-experimental studies that can identify causal effects. We identify 11 papers published between 2009 and 2019, evaluating 23 policy changes across Europe and North America from 1977 to 2009. Results are a mixture of positive, negative, and null impacts on fertility. To explain these apparent inconsistencies, we extend the conceptual framework of Lalive and Zweimüller (2009), which decomposes the total effect of leave on fertility into the "current-child" and "future-child" effects. We decompose these into effects on women at different birth orders, and specify types of study design to identify each effect. We classify the 23 studies in terms of the type of effect identified, revealing that all the negative or null studies identify the current-child effect, and all the positive studies identify the future-child or total effect. Since the future-child and total effects are more important for promoting aggregate fertility, our findings show that leave does in fact increase fertility when benefit increases are generous. Furthermore, our extensions to Lalive and Zweimüller's conceptual framework provide a more sophisticated way of understanding and classifying the effects of pro-natal policies on fertility. Additionally, we propose ways to adapt the ROBINS-I tool for evaluating risk of bias in pro-natal policy studies.

Introduction

Persistent low fertility can cause economic problems, and is associated with parents having fewer children than they want. Many countries in East Asia and Europe have fertility levels significantly below the replacement level of 2.1 births per woman, and countries across Asia and Latin America are projected to join them in the coming decades (UN DESA, 2019). When low fertility rates persist, rising old-age dependency ratios and reductions in the working-age population are theorised to cause higher consumption, lower investment, and economic

---

[1] University of Liverpool, Liverpool, UK. [2] National Tsing-Hua University, Hsinchu, Taiwan (R.O.C.). ✉email: jacthom@liverpool.ac.uk

stagnation or decline (Bloom et al., 2010; Caldwell, Caldwell and McDonald, 2002; McDonald, 2008). On the individual level, low fertility can be indicative of people having fewer children than they would ideally like to have (Beaujouan and Berghammer, 2019; Chen and Yip, 2017; Spéder and Kapitány, 2014). The failure of individuals to fulfil their childbearing intentions can negatively impact emotional well-being (Casterline and Han, 2017; Priebe, 2020; Ugur, 2020).

Governments in low-fertility countries have used family policies to promote fertility (Gauthier, 2007; Raute, 2019; Rindfuss and Choe, 2016). Family policies typically aim to support parents in their caring responsibilities, and enhance parents' and children's well-being (Eydal and Rostgaard, 2018; Gauthier, 2008). However, increasing fertility is rarely an explicit objective of family policies, and is regarded more as a potential by-product of those policies (Thévenon, 2011). Family policies can increase fertility either through helping parents balance work and family, or through reducing the costs of childbearing and childrearing (Gauthier, 2007; Gauthier and Philipov, 2008; Rindfuss and Choe, 2016). Family policies can be categorised as child-related cash transfers, childcare subsidies, or financial support through the tax system (OECD, 2019a). Child-related cash transfers tend to take up the largest proportion of public expenditure, amounting to 1.3% of GDP across OECD countries in 2015, a total spend of over $0.75 trillion (OECD, 2019a, 2019b, 2019c). Within child-related cash transfers, parental, maternity and paternity leave policies (henceforth 'leave') refer to state mandated arrangements for parents to take time off work during pregnancy or after childbirth. In 2015, average OECD public expenditure on leave was $12,100 per infant (at purchasing power parity, 2010 USD) (OECD, 2019d). Understanding the extent of impacts of leave on fertility is thus critical to for governments aiming to increase fertility through family policies.

This paper focusses on leave—rather than other family policies —because leave increases domestic gender equity and domestic gender inequity has been identified as a key cause of low fertility (Goldscheider et al., 2015; McDonald, 2006; Tamm, 2019). Over the past two decades, researchers have increasingly argued that the tension between increasing career ambitions and persistently gendered domestic obligations has meant that women's ability to reconcile work with family life has become more restricted, which in turn seems to have contributed to low fertility (e.g., Baizan et al., 2016; Duvander, Johansson and Lappegard, 2016; Meier and Rainer, 2017). Maternity, paternity and parental leave can help equalise the division of domestic and formal labour between men and women through enabling mother's return to work, and through encouraging fathers to do more housework and childcare[1] (Baum and Ruhm, 2016; Pronzato, 2009; Tamm, 2019). Through equalising the gender balance of labour, leave can reduce the cost of childrearing for women, and thereby increase fertility (Baizan et al., 2016; Kotsadam and Finseraas, 2011). Although gender equity is our motivation for focussing on leave, we only seek to establish whether leave has an impact on fertility.

Whether leave actually increases fertility remains a matter of debate (Balbo et al., 2013; Hoem, 2008; Olivetti and Petrongolo, 2017). Gauthier's (2007) review of the effect of family policies on fertility found the evidence to be "mixed," an evaluation echoed across the empirical literature (e.g., Hong and Sullivan, 2016; Lappegård, 2010; Matysiak and Szalma, 2014). Similarly, Bergsvik, Fauske and Hart's review of the effect of family policies on fertility (2020) concludes that the effect of leave on fertility is ambiguous. There are certainly cases in which governments have provided generous leave policies and low fertility has persisted, such as in Slovenia (Stropnik and Sircelj, 2008). There are also cases in which generous new leave policies have been accompanied by large increases in fertility (e.g., East Germany), and

cases in which generous leave policies have been accompanied by stable and high fertility (e.g., Czechoslovakia and Sweden) (Buttner and Lutz, 1990; Hoem, 1990, 1993, 2005; Monnier, 1990; Salles, 2006). However, most empirical studies use methods that prohibit identification of a causal effect of leave on fertility. To date, there have been no peer-reviewed systematic reviews focussing on leave and fertility, which discriminate between studies that can identify causal effects and studies that cannot. Whether leave does in fact cause higher fertility, therefore, remains an open question. A peer-reviewed systematic review could reliably answer this question by showing whether or not more generous parental leave leads to higher fertility, thereby resolving existing academic debates and giving governments a sounder footing for policymaking.

This paper provides the first peer-reviewed systematic review of the effect of maternity, paternity, and parental leave on fertility. In this systematic review we seek to evaluate the effects of leave policies on fertility, in order to inform policymakers who are considering leave as a means of increasing fertility. We aim to find, evaluate, and synthesise all relevant, experimental or quasi-experimental studies, in a rigorous, transparent, and reproducible fashion. We seek to answer the question: to what extent does leave increase fertility? Since we are interested in informing pro-natal policies in low-fertility countries, we restrict our search to countries and time periods where fertility is persistently below 2.1 births per woman (broadly speaking, in high-income countries from the 1950s onwards). Using a thorough search of all published English-language material catalogued online, we identify 11 papers that match our inclusion criteria, containing 23 different studies that can plausibly test for a causal effect of leave on fertility.[2]

Our paper makes three contributions. Firstly, we extend the conceptual framework of Lalive and Zweimüller (2009) (section "Effects of leave on fertility: conceptualisation and identification"), building on their definitions of the "current-child effect" and the "future-child effect". We extend their framework by decomposing the impact of leave on individuals by type of effect and by parity, by specifying study designs that can be used to identify each impact, and by exploring, which effects are most important for policymakers aiming to increase fertility. This extended framework can also be applied to assess the impact of other pro-natal policies on fertility. Secondly, we find that paternity or parental leave reforms, which provide generous increases in duration or remuneration consistently increase fertility, implying that large increases in such benefits are a viable strategy for governments seeking to raise fertility (section "Results"). Maternity leave reforms are also found to increase fertility, although may act to decrease fertility by reinforcing traditional gender norms. Categorising the 23 studies in terms of our framework reveals that studies whose methods identify a broader class of effects consistently find positive results, and that all negative or null studies only address a narrow class of effects that are of marginal interest to policymakers. Thirdly, we propose ways to adapt the ROBINS-I tool for the evaluation of studies of pro-natal policies (section "Study quality"). ROBINS-I was designed as a tool for assessing risk of bias (RoB) in non-randomised studies of medical interventions (Sterne et al., 2016); we identify three key reasons why it is not directly applicable to studies of public policy interventions. First, ROBINS-I assumes the existence of placebo effects, which do not exist in a public policy context. Second, the notion of an "intention-to-treat" (ITT) study is problematic in the case of public policy because individuals can self-select into being (in)eligible for a policy after their initial assignment (in medical studies, participants cannot change their assignment status after being assigned or not assigned to treatment). Third, ROBINS-I does not distinguish

between policy eligibility and policy availability, which comprise critical dimensions of leave policies. As well as identifying these three considerations and proposing strategies to account for them, we also identify the sources of bias inherent in each of the study designs in our conceptual framework, providing future empirical researchers with a checklist of sources of bias and strategies to minimise RoB.

The paper is structured as follows. Section "Background" provides background information on the mechanisms by which leave is theorised to affect fertility, and introduces Lalive and Zweimüller's (2009) definitions of the current-child and future-child effects. Section "Methods" describes systematic review methods and explains how we conducted this review. Section "Results" presents the results of our search and filtering process, and develops our conceptual framework to classify and assess studies and the type of fertility effect they capture. Section "Results" also assesses the RoB of included studies, and synthesises the study findings. Section "Discussion" discusses our analysis and findings, and Section "Conclusion" offers some final concluding remarks.

Background

This section provides background information on three topics. Section "Reasons for introducing or extending leave" explores the various justifications for introducing or extending leave, Section "Theory of how leave affects fertility" describes the various channels by which leave is theorised to increase fertility, and Section "Effects on fertility: current-child effect and future-child effect" provides an exposition of Lalive and Zweimüller's (2009) theory of the current-child effect and the future-child effect.

Reasons for introducing or extending leave. The first leave policies were maternity leave policies, which were introduced in the early twentieth century for two reasons: firstly, to give mothers the right to return to work after a period of absence; and secondly, to protect the health of infants and mothers (Gauthier and Bartova, 2018; Gauthier and Koops, 2018; Zabel, 2009). In the contemporary world, leave policies have a broader set of aims, including work-family reconciliation and gender equity (ibid.). Researchers have identified six main objectives of family policies: poverty reduction and income maintenance; direct compensation for the economic costs of children; fostering employment; improving gender equity; supporting early childhood development; and increasing fertility (Nieuwenhuis and Lancker, 2020; Thévenon, 2011). All six objectives are applicable to leave policies, with the possible exception of poverty reduction. Increasing fertility has rarely been an explicit goal of leave policy (Gauthier and Bartova, 2018; Thévenon, 2011); however, governments in low-fertility countries—such as China and Germany—are increasingly turning to leave policies as a way of increasing fertility (Spiess and Wrohlich, 2008; Raute, 2019; Wanqing, 2021). Regardless of politicians' stated justifications for introducing or extending leave, all leave policies have the potential to reconcile childbearing and employment, and thereby increase fertility. In this systematic review we seek to evaluate the effects of leave policies on fertility, in order to inform policymakers who are considering leave as a means of increasing fertility. We are therefore indifferent to the various other stated aims of leave, and are only interested in the impact of leave on childbearing.

Theory of how leave affects fertility. Leave is theorised to increase micro-level fertility by enabling couples and individuals to realise their childbearing aspirations, through facilitating a balance between work and childcare and by lowering the net costs of childrearing (Gauthier, 2008). However, the impact of leave on fertility is complicated by the potentially ambiguous effect of leave policies on gender roles, and how that effect is mediated by the type of welfare regime of the country in question. There is a large body of research on international leave policies and their impacts—for reviews, see the annual reviews of the International Network on Leave Policies and Research (https://www.leavenetwork.org/), Gauthier and Bartova (2018), and Hegewisch and Gornick (2011). Here, we focus on the impacts of leave on the costs of childbearing and childrearing, and how these impacts are mediated by gender roles and welfare regimes. The causal mechanisms by which leave is theorised to impact fertility are presented in a numbered list, with five mechanisms identified. We are interested in the impact of the availability of leave on couples' decisions to have a child, rather than the impact of leave use on couples' decisions to have a child; this distinction is explored further in section "Effects on fertility: current-child effect and future-child effect".

Work-life balance and financial costs. In the absence of leave, an employed parent can either return to work soon after childbirth, or quit work in order to look after the child. Leave enables parents to take some partially remunerated weeks or months off, and then return to work. While parents may lose some income and not progress in their career while on leave, they are able to spend time with their infant. For individuals who want to maintain their career but also have children, leave enables them to balance these two preferences. Consequently, individuals with such preferences might have children they would not have had in the absence of leave (Becker, 1973; Ermisch, 2003). This mechanism is particularly important for maternity leave, because women are expected to be primary care-givers.

(1) Leave facilitates childbearing for parents who prefer to balance caring for their infant with maintaining their current job.
Regardless of employment preferences, taking leave might have the lowest net financial cost: if the rate of remuneration of leave is high and childcare is expensive, then taking leave will be cheaper than either other option (Gauthier, 2008). Leave can therefore lower the net costs of childrearing, facilitating childbearing:
(2) Leave may lower the net financial costs of rearing infants, facilitating childbearing.

Mechanism (2) is relevant for both maternity leave and paternity leave.

Gender. The gender balance of domestic and formal labour in a country, and how leave policies interact with that balance, are important determinants of fertility (Goldscheider et al., 2015; McDonald, 2000; Raybould and Sear, 2020). Some researchers have argued that leave can encourage more egalitarian roles and behaviour, whereas others have argued that leave can reinforce traditional gender roles and behaviour; for reviews, see Farré (2016), and Hegewisch and Gornick (2011). Theoretically, leave policies that promote gender equity should increase fertility, and the converse for policies that promote inequity (McDonald, 2000; McDonald, 2006).

On the positive side, leave may increase fertility through promoting gender equity. McDonald's "gender equity theory" argues that fertility in developed countries will be low where family-oriented institutions assume a male-breadwinner model of the family, and public institutions (such as education and employment) assume a gender-equal model, since together they represent incoherent expectations or prescriptions for women's domestic and work lives (McDonald, 2000; McDonald, 2006). Gender inequity therefore reduces fertility by constraining the

choices of women. Leave can increase fertility by increasing equity, either via maternity leave enabling mothers to remain in the labour market (Geyer et al., 2015), or paternity leave encouraging fathers to participate more in housework and childcare, equalising the gendered division of domestic labour (Farré, 2016):

(3)   Maternity leave enables mothers to remain in employment, reducing gender inequity and thereby increasing subsequent fertility.

(4)   Paternity leave promotes men doing domestic labour and enables mothers to work, reducing gender inequity and thereby increasing subsequent fertility.
       On the negative side, maternity leave may decrease gender equity by reinforcing traditional gender roles. Most countries provide longer periods of leave to mothers than fathers: out of 15 EU countries in 2002, only Sweden was starting to develop gender-equal leave entitlements (Haas, 2003). Therefore, in terms of entitlements, most existing leave systems assume a model of the family in which women are to some extent the primary care-givers, thereby possibly "trapping" mothers at home (Evertsson and Duvander, 2011). In line with gender equity theory, maternity leave may therefore reduce fertility:

(5)   Maternity leave may reinforce traditional gender roles, decreasing gender equity and thereby decreasing subsequent fertility.

In terms of evidence, several articles review the effects of leave policies on outcomes such as domestic gender equity, the motherhood wage penalty, and women's labour force participation (e.g., Dearing, 2015; Gauthier and Bartova, 2018; Hegewisch and Gornick, 2011). These reviews highlight that leave positively impacts female labour force participation, though longer periods of leave are associated with larger motherhood wage penalties. Dearing's review (2015) finds the evidence for paternity leave promoting fathers' domestic work to be mixed; however, empirical articles not included in that review seem to support a positive relationship more consistently, in countries such as Germany and Spain (Bünning, 2015; Fernández-Cornejo et al., 2018; Schober, 2014). Overall, the evidence seems to suggest that leave promotes a more equal balance of domestic and formal labour between men and women, and especially when maternity leave is not too long.

Welfare regimes. Within a given country, the impact of leave on fertility is influenced by the way in which state welfare provisions are gendered. As well as for other social institutions more broadly, different welfare regimes are based around different models of the family, such as a male-breadwinner model or gender-egalitarian model (McDonald, 2000). These assumed models of the family influence the ways in which welfare regimes set family and labour policies, including the extent to which they provide alternative childcare arrangements for mothers wanting to balance work with childrearing (Baxter et al., 2008; Budig et al., 2015; Neilson and Stanfors, 2014). Since Esping-Andersen's (1990) seminal welfare regime typology, newer typologies have sought to incorporate the role of gender relations and family policies (e.g., Ciccia and Verloo, 2012; Saxonberg, 2013). Saxonberg (2013) evaluates welfare regimes both in terms of the benefit levels of parental leave and in terms of state support for day care, and classifies regimes into degenderizing, explicitly genderizing, and implicitly genderizing types. Ciccia and Verloo (2012) use fuzzy-set matching to quantify countries' adherence to six different ideal types, ranging from a gender-equitable full universal caregiver (FUC) type, to a gender-inequitable male-breadwinner (MB) type. Broadly, both typologies tend to classify Scandinavian

and Benelux countries as promoting gender-equitable outcomes, and Southern, Central and Eastern European countries, as well as Anglophone countries, as reinforcing traditional norms. The gendered configuration of welfare regimes acts to either promote or impede mothers shouldering the majority of childcare and domestic labour, which complicates the identified causal mechanisms (1)-(5). For example, gender-inequitable regimes may have higher motherhood earnings penalties, mitigating causal mechanism (2) (Budig et al., 2015; Fernández-Cornejo et al., 2018). In this review, we therefore consider welfare regimes as influencing the causal impacts of leave on fertility, rather than acting as an independent causal agent on fertility.

A benefit of systematic reviews such as this is that included studies are limited to quasi-experimental designs, which filter out the potentially confounding role of national contexts. This means that included studies are able to reliably identify the effects of leave policy changes, though the size of those effects will still be influenced by the welfare regime in question.

Effects on fertility: current-child effect and future-child effect. Leave can affect fertility behaviour in two key ways. In a seminal analysis in the econometrics literature, Lalive and Zweimüller (2009) distinguish between the "current-child effect" and the "future-child effect" of a leave policy on fertility. This conceptual framework has become commonly used in other studies, in order to elucidate which type of effect is being identified (e.g., Dahl et al., 2016; Cygan-Rehm, 2016; Raute, 2019).

The current-child effect refers to the effect of being able to take more leave for the child just born, on subsequent fertility. Women giving birth shortly before and shortly after a reform receive different benefits for the child they just had. However, both groups of women will receive equal benefits for any subsequent children. Therefore, if there is any long-term difference in fertility between the two groups, it must be due to the different benefits they received for the child born around the time of the reform.

The future-child effect refers to the effect of a greater amount of leave available in the future. It is called the "future" child effect to distinguish it from the current-child effect, and captures the idea that, if a woman knows she will receive more generous leave entitlements if she has a child, she will probably be likelier to have that child. Lalive and Zweimüller identify the future-child effect by comparing the fertility of mothers in the years before the reform, with the fertility of mothers in the years after the reform.

Lalive and Zweimüller go on to argue that the sum of the current-child effect and the future-child effect gives the total effect of leave on fertility. However, we note that their study design can only identify effects on women at parities of 1 or higher, since the women in their samples all had at least one birth. This means that their study design cannot identify the effect of leave on women with no children, and therefore cannot identify the total effect of leave on fertility.[3] While other studies have used Lalive and Zweimüller's terminology, no study has highlighted the role of parity in classifying effects. Separating the impact of a leave policy on women at different parities can enable us to understand the different processes by which individuals choose to have a child.

Methods
Our methods were guided by the policies and guidelines of The Campbell Collaboration for conducting systematic reviews (https://campbellcollaboration.org/) (Campbell Collaboration, 2019). Prior to conducting literature searches we produced a review protocol, which was registered and published online at the International prospective register of systematic reviews (PROSPERO), on the 9th of September 2019 (Thomas et al., 2019). Full

details of our method are given in the review protocol, which is available online at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019128493. This report was written according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Moher et al., 2009). Supplementary material to this paper is provided online, and all appendices referenced here appear in that supplementary material. All deviations from the protocol are specified in Supplementary Appendix A.

Criteria for inclusion of studies. Studies must be primary, empirical, quantitative studies that assess the effect of a leave policy on fertility at the micro-level. We exclude macro-level studies since they cannot identify the causal mechanisms by which policy affects fertility (Neyer, and Andersson, 2008). We are interested in the effects of policy changes involving one or more of maternity, paternity, and parental leave. Changes can either be in increasing, decreasing, or restructuring leave. Since the state is the policymaker of interest, we are not interested in the effects of firm-specific policies on fertility. Furthermore, since we aim to collect evidence to inform policy in low-fertility countries, we are only interested in finding studies of policies implemented in countries with a TFR below 2. We apply this restriction because we want the included studies to have external validity in terms of taking place in settings that are comparable to other countries with low fertility (Shadish et al., 2002). Countries and time periods that are eligible for inclusion are specified in the protocol (Thomas et al., 2019). In terms of fertility, we are interested primarily in quantum effects rather than tempo effects. We exclude articles that purely consider policy effects on birth timing or seasonality.

We only include studies with strong designs that can plausibly provide evidence of a causal relationship between leave changes and fertility. These study designs are: randomised control trials (RCTs), quasi-experimental designs, and natural experiments, which can all provide evidence of a causal relationship through specifying appropriate means of estimating counterfactual situations. These designs therefore have a higher degree of internal validity than purely correlational or observational studies (Shadish et al., 2002). A key element in identifying a cause is establishing that the cause occurred before the effect, and so we are interested only in the effects of leave policy *changes* on fertility, rather than any relationship between existing leave policies and fertility.

Finally, we only consider studies that evaluate the effect of leave availability on fertility, rather than studies that evaluate the effects of leave uptake or use on fertility. This decision is motivated by the Cochrane Collaboration's advice that evaluating "intention-to-treat" (ITT) effects tends to result in less biased outcomes than "per-protocol" effects (Higgins and Green, 2011). Since parents self-select into using leave, parents who use leave are likely to differ systematically from those who do not, meaning that the causal effect of leave on fertility cannot be identified (Lappegård, 2010). Moreover, the availability of leave can motivate parents to conceive even if they do not take leave after childbirth. This means that only examining leave use cannot capture the full effect of the policy on fertility, which is of most interest to policymakers looking to increase fertility.

Search strategy and filtering process. A flowchart illustrating our search and filtering process is given in Fig. 1. The search process was divided into searching academic databases, searching grey literature sources, hand-searching relevant journals, and snowball searching using the references and citations of included articles. Filtering of included articles was done on the basis of titles, then

abstracts, then full texts. Supplementary Appendix B details the search procedure for one database, "Academic Search Complete."

Assessing study quality. To assess study quality we used the ROBINS-I tool for assessing Risk of Bias (RoB) in non-randomised studies (Sterne et al., 2016). ROBINS-I is applied separately to each study in the review, and works by comparing the (non-randomised) study to a hypothetical, idealised RCT, in which there would be no RoB. Using ROBINS-I involves answering a series of signalling questions on characteristics of each study, across seven domains of bias: confounding, selection bias, misclassification bias, performance bias, bias from missing data, detection bias, and outcome reporting bias (these domains are explained more fully in Supplementary Appendix C). The answers to the signalling questions are used to generate an overall classification of the study's RoB, ranging from "low risk of bias" (where the study is considered to be comparable to a well-performed RCT), to "critical risk of bias" (where the study is too problematic to provide any useful evidence, and should not be included in the synthesis). The signalling questions for each domain are determined by whether the study aims to measure the effect of "assignment to intervention," or the effect of "assignment and adherence to intervention."

Method of synthesis. Owing to the many dimensions of difference between leave policies, it would be inappropriate to attempt any kind of statistical meta-analysis of the study results. Instead, we conducted a narrative review of included studies, using the Economic and Social Research Council's (ESRC) "Guidance on the Conduct of Narrative Synthesis in Systematic Reviews," document (Popay et al., 2006).

Results

This section provides an overview of the 11 papers and 23 studies from our search, and shows their results to be mixed. However, we reveal that there is a key underlying reason for the apparently mixed evidence. To this end, we extend Lalive and Zweimüller's (2009) conceptual framework in order to understand the different fertility effects identified by the studies. We use this framework to classify the 23 studies, and then evaluate their RoB. Finally we synthesise the results.

Overview. The search and filtering process—with the number of articles removed at each stage—is summarised in Fig. 1. We separately searched for articles in both the academic and grey literature. We first scanned the academic literature, returning 5470 results; after removing duplicates, 2996 remained. Filtering based on title and abstract reduced these to 51 papers. After careful reading of the full text, this sample was further reduced to 7 papers. Secondly, we swept the grey literature returning 528 results. Filtering based on title and abstract reduced these to 3 results, and after reading the full texts 2 articles were preserved for the final analysis. We also conduced hand searches of relevant journals, but this returned no new results. We then conducted a snowball search of references and citations, which returned 2 further results. Thus, 11 was the final number of articles included for analysis.

Table 1 summarises key information on the 11 papers and 23 studies, including author names and publication year, year of the reform, country, type of leave, sample size, dependent variable, and effect found. The focus of the studies varies slightly. Two studies evaluated maternity leave; one study evaluated paternity leave; and 19 studies evaluated parental leave—although these policy changes differed in whether they affected mothers or fathers in practice. One study evaluates a reform that
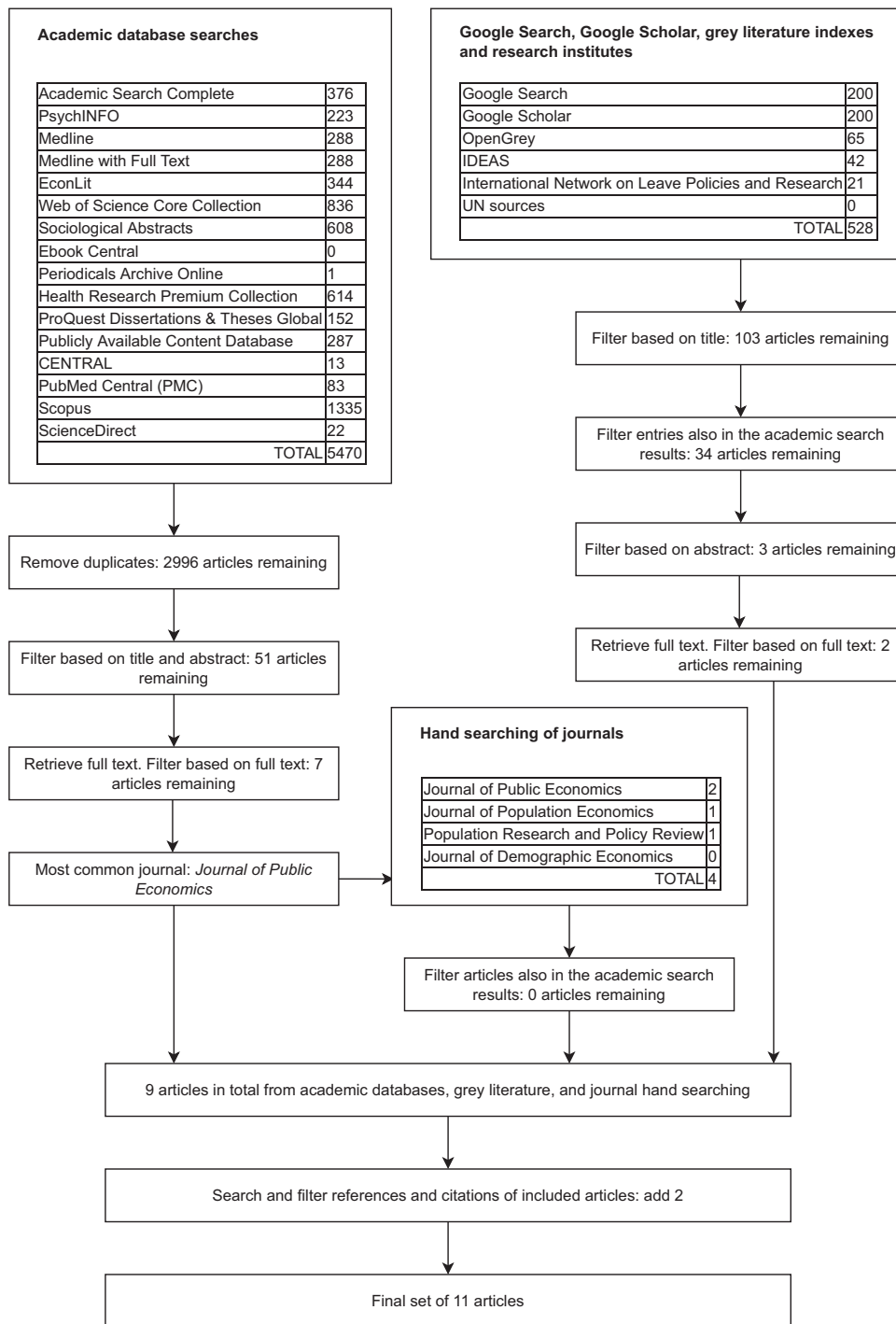
Academic database searches

Academic Search Complete	376
PsychINFO	223
Medline	288
Medline with Full Text	288
EconLit	344
Web of Science Core Collection	836
Sociological Abstracts	608
Ebook Central	0
Periodicals Archive Online	1
Health Research Premium Collection	614
ProQuest Dissertations & Theses Global	152
Publicly Available Content Database	287
CENTRAL	13
PubMed Central (PMC)	83
Scopus	1335
ScienceDirect	22
TOTAL	5470

Google Search, Google Scholar, grey literature indexes and research institutes

Google Search	200
Google Scholar	200
OpenGrey	65
IDEAS	42
International Network on Leave Policies and Research	21
UN sources	0
TOTAL	528

Filter based on title: 103 articles remaining

Filter entries also in the academic search results: 34 articles remaining

Filter based on abstract: 3 articles remaining

Retrieve full text. Filter based on full text: 2 articles remaining

Remove duplicates: 2996 articles remaining

Filter based on title and abstract: 51 articles remaining

Retrieve full text. Filter based on full text: 7 articles remaining

Most common journal: *Journal of Public Economics*

Hand searching of journals

Journal of Public Economics	2
Journal of Population Economics	1
Population Research and Policy Review	1
Journal of Demographic Economics	0
TOTAL	4

Filter articles also in the academic search results: 0 articles remaining

9 articles in total from academic databases, grey literature, and journal hand searching

Search and filter references and citations of included articles: add 2

Final set of 11 articles

Fig. 1 The literature searching and filtering process. Boxes indicate the sources used and filtering processes.

included all three of parental, paternity and maternity leave. The studies cover 7 countries across Western Europe and North America. The vast majority of studies evaluate reforms implemented in the 1990s and 2000s, with only three studies evaluating reforms before 1990. The length of follow-up varies widely, from 2 years to 30 years.

Eight studies found a positive effect of leave on fertility, one found a negative effect, and 14 found no evidence of an effect. It therefore would seem like the evidence on the effect of leave on fertility is mixed. However, Table 1 does not indicate which type of effect—current-child or future-child—each study identified. The reason for this is that some of the included studies measure

effects that cannot be identified in Lalive and Zweimüller's original classification.

Effects of leave on fertility: conceptualisation and identification. We now extend Lalive and Zweimüller (2009) to develop a formal conceptual framework to assess the effect of leave on fertility. Section "Types of effects of leave on fertility" extends their terminology to include effects on individuals at all parities, Section "Study designs and identification strategies" explains the empirical strategies used to identify each effect, and Section "Categorisation of the studies in the review" classifies the 23 studies in terms of our framework.

Table 1 Key characteristics of included studies.

Author(s) (year of publication)	Year of reform	Country	Type of leave	Sample size	Dependent variable	Effect of the new policy (interpretation)
Ang (2015)	2006	Canada	Maternity, paternity, and parental	2,048,800	Child born in Quebec up to 2008	0.012*** (women 23.53% likelier to have a birth)
Cannonier (2014)	1993	USA	Maternity	11,963	First child born in the next 17 years	0.0519** (women 5.19 percentage points likelier to have a birth)
				22,263	Second child born in the next 17 years	0.0296* (women 2.96 percentage points likelier to have a birth)
Carneiro, Løken, and Salvanes (2015)	1977	Norway	Parental (maternity in practice)	63,571	Number of children born in next 30 years	−0.011 (women had 0.011 less children)
Cools, Fiva, and Kirkebøen (2015)	1993	Norway	Parental (paternity in practice)	23,985	Child born in the next 14 years	0.004 (fathers had 0.004 more children)
Cygan-Rehm (2016)	2007	Germany	Parental	39,826	Child born in next 5 years	0.046 (women 4.6 percentage points likelier to have a birth)
Dahl et al. (2016)	1987	Norway	Parental (maternity in practice)	17,580	Number of children born in the 14 years after each reform	−0.030 (women had 0.030 fewer children)
	1988			19,310		0.017 (women had 0.017 more children)
	1989			20,093		−0.016 (women had 0.016 fewer children)
	1990			21,508		−0.026 (women had 0.026 fewer children)
	1991			21,717		−0.024 (women had 0.024 fewer children)
	1992			21,838		0.042* (women had 0.042 more children)
Duvander, Johansson, and Lappegard (2016)	1993	Norway	Parental (paternity in practice)	1804	Second child born in the next 10 years	0.98 (couples 2 percentage points less likely to have a birth)
				1702	Third child born in the next 10 years	0.94 (couples 6 percentage points less likely to have a birth)
	1995	Sweden	Parental (paternity in practice)	3880	Second child born in the next 10 years	0.95 (couples 5 percentage points less likely to have a birth)
				3383	Third child born in the next 10 years	0.94 (couples 6 percentage points less likely to have a birth)
Farré and González (2019)	2007	Spain	Paternity	(Not reported)	Child born in the next 6 years	−0.0179*, though not robust to bandwidth (fathers 5.3% less likely to have a birth)
Hart, Andersen and Drange (2019)	2009	Norway	Parental (paternity in practice)	9757	Child born in the next 5 years	−0.04 (couples had 0.04 fewer children)
	1990	Austria		6180		

Table 1 (continued)

Author(s) (year of publication)	Year of reform	Country	Type of leave	Sample size	Dependent variable	Effect of the new policy (interpretation)
Lalive and Zweimüller (2009)			Parental (maternity in practice)	5997	Child born in the next 10 years	0.035** (women 3.5 percentage points likelier to have a birth)
					Child born in the next 3 years	0.068** (women 6.8 percentage points likelier to have a birth)
	1996		Parental (maternity in practice)	(Not reported)	Child born in the next 3 years	−0.001 (women 0.1 percentage points less likely to have a birth)
				(Not reported)	Child born in the next 3 years	Positive effect (numerical result not reported)
Raute (2019)	2007	Germany	Parental	644,981	Child born to affected women after the reform	6.424*** (women 16% likelier to have a birth)

For the numerical results reported in the "Effect of the new policy" column, * indicates $p < 0.1$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$.

Types of effects of leave on fertility. Lalive and Zweimüller argue that the total effect of a leave policy on fertility is the sum of the current-child effect and the future-child effect. However, they could not identify the effect of the policy on women at parity 0. We use the term "future-child effect (parity 1+)" in the sense that Lalive and Zweimüller use "future-child effect," to mean an effect on women at parities of 1 or higher. In order to explain studies that evaluate the effect on women at parity 0, we introduce the term "future-child effect (parity 0)" to mean the effect of the policy on women who have not had any children. We use the term "total effect (parity 1+)" to mean the total effect identified by Lalive and Zweimüller, and "total effect" to mean the sum of all these effects across the population, as displayed in Fig. 2 below.

As discussed in section "Reasons for introducing or extending leave", an aim of this systematic review is to inform policymakers in low-fertility countries who may be considering extending leave as a way to increase fertility. We contend that such pro-natal policymakers are more interested in the future-child effect and the total effect, than they are in the current-child effect. In the language of experimental design, the future-child effect and the total effect have high "construct validity," and the current-child effect has low construct validity (Shadish et al., 2002). "Construct validity" refers to whether the specific features of an experiment validly capture the underlying concepts, or "target constructs" (ibid.). When designing a leave policy for pro-natal purposes, policymakers have some concept of "the effect of leave on fertility." We contend that policymakers have one of two conceptualisations of this concept: either "the effect of the availability of leave—for a yet unborn child—on the decision of an individual to have that child," or "the overall effect of leave on the fertility of women across the population." These two conceptualisations clearly correspond to the future-child effect and the total effect. We therefore judge studies identifying the future-child effect or the total effect to have high construct validity, because these studies validly represent the target construct of "the effect of leave on fertility." By contrast, the current-child effect would correspond to the conceptualisation, "the effect of leave on someone who has just had a child, on their subsequent childbearing." We contend that this conceptualisation is not what pro-natal policymakers mean by "the effect of leave on fertility," and so we judge the current-child effect to have low construct validity. For the remainder of this review, we will divide studies into two categories: current-child effect, and future-child and total effect. Since our objective is to inform pro-natal policymaking, we will give greater weight to future-child and total effect studies.

Study designs and identification strategies. The current-child effect can be identified by comparing those giving birth in the weeks or months before the reform, with those giving birth in the weeks or months after the reform. Since these two groups differ in terms of their leave entitlements for the child they just had—and will have the same entitlements for any future child —the current-child effect can be identified by comparing their subsequent fertility over the next several years. Such a study design attempts to approximate a randomised study by arguing that women giving birth shortly before and after the reform are likely to be otherwise similar, meaning that other variables are controlled for. Such a study design is a type of regression discontinuity design (RDD). We define this study design as the "short before-after" design, and illustrate how it works in Fig. 3(a).

The short before-after design typically requires large administrative datasets, since survey data typically will only have very few women giving birth in the short periods before and after the reform. Since both groups have access to the new policy after the reform, the short before-after design cannot identify the future-
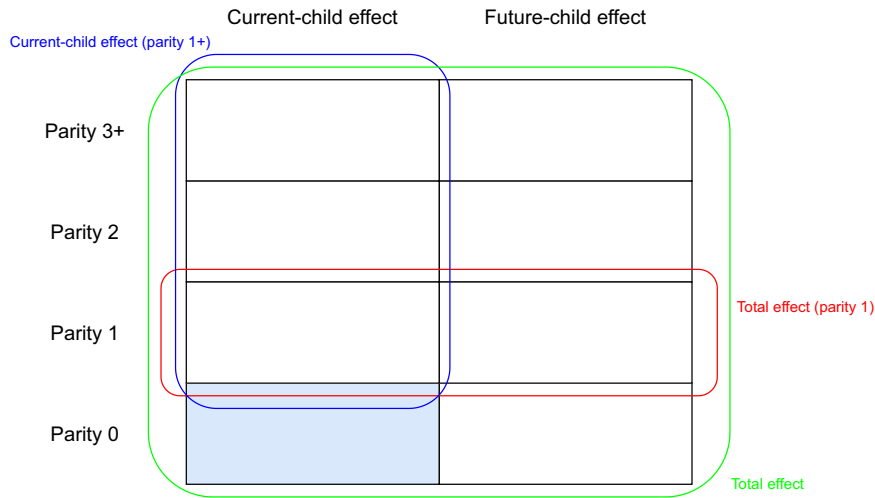
Fig. 2 Types of effect of a leave policy on fertility, by parity. The two columns correspond to the current-child effect and the future-child effect. The rows correspond to the parity of the parent. The blue, red, and green oblongs represent the current-child effect (parity 1+), the total effect (parity 1), and the total effect, respectively.
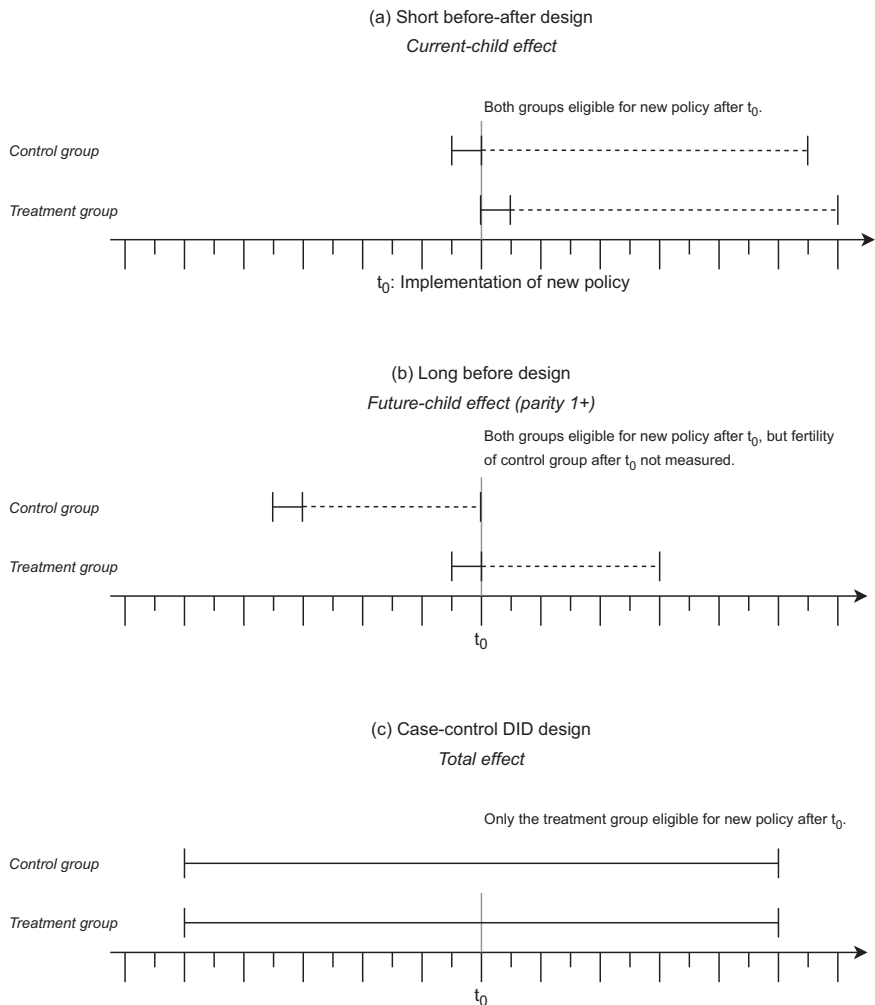

Fig. 3 Study designs for evaluating the effects of leave on fertility. a The short before-after study design, used to evaluate the current-child effect. **b** The long before study design, used to evaluate the future-child effect on individuals of parity 1 or higher. **c** The case-control difference-in-differences study design, used to evaluate the total effect. Long check marks on the horizontal axis indicate years, and short check marks indicate half-years. Solid horizontal lines indicate data on births being collected at that time, and dashed horizontal lines indicate that data on births occurring at that time will be collected later.

child effect (parity 1+). Furthermore, the short before-after design cannot identify the future-child effect (parity 0), since the analytical sample is restricted to women who have had at least one child.

To identify the future-child effect (for any parity), an alternative study design is required. Lalive and Zweimüller aim to identify the future-child effect (parity 1+) by comparing women who gave birth a month before the reform, with women who gave birth in the same month 3 years before. Both cohorts are therefore entitled to the old policy for the child they have just had, but will have different entitlements if they have another child within the next 3 years. The future-child effect (parity 1+) can then be identified by comparing the fertility of the two cohorts at the end of their respective three year periods. We define this design as the "long before" study design, and illustrate it in Fig. 3(b). A major problem with the long before identification strategy is that it assumes there are no systematic differences (that are important for fertility) between the two cohorts. This is a strong assumption: in order for it to be valid, there would have to be no effect of long-term trends in childbearing behaviour. Consequently, estimates for the future-child effect (parity 1+) under this strategy are at risk of bias.

The long before design cannot identify the future-child effect (parity 0) for the same reason that it cannot be identified in the short before-after design: in order to be included in the analytical sample, individuals must have had at least one child. Consequently, neither design can estimate the total effect. In order to estimate the total effect, a case-control study design is required. This means that the new leave policy only becomes available to some women (the case group, who receive the "treatment" of the policy) and not others (the control group, who do not receive treatment). The total effect can then be estimated by comparing the post-reform fertility of the case group and the control group. In a randomised study, individuals would be assigned to treatment randomly; however, this is rarely the case with policy changes, and not the case for any of the studies in this review. In the three case-control studies included in this review, individuals are allocated to treatment by either region, employment status, or income. These studies are problematic in the sense that there are usually pre-existing systematic differences between the case group and the control group, differences, which are correlated with either region, employment status, or income. However, a difference-in-differences (DID) approach enables one to approximately control for these differences. We therefore use the term "case-control DID" to refer to this type of study design. A possible implementation of the case-control DID design is illustrated in Fig. 3(c), using panel data. DID models make two key assumptions: the common trends assumption (CT), and the common support assumption (COSU) (Angrist and Pischke, 2009; Keng and Sheu, 2011; Lechner, 2010). CT requires trends in the case and control groups to be roughly parallel prior to the introduction of the policy, and COSU requires that the distributions of other predictors of the outcome variable must remain roughly similar over time (Lechner, 2010). Whether these assumptions are valid in each of the studies is evaluated in the RoB section.

While it is possible to identify the total effect using a case-control DID design, the precise effect identified in any given model depend on the sample restrictions imposed by the analyst. For example, in one of the case-control DID models used by Cannonier (2014), the analytical sample is restricted to women who had never given birth at time $t_0 - 4$. The outcome variable is a binary indicator, indicating whether the individual had at least one birth between $t_0 - 4$ and $t_0 + 17$. This method therefore identifies the future-child effect (parity 0). By contrast, in Cannonier's second model, the analytical sample is restricted to

women who had had exactly one birth at time $t_0 - 4$. The outcome variable is a binary indicator, indicating whether the individual had at least one more birth between $t_0 - 4$ and $t_0 + 17$, and so this model identifies the total effect (parity 1).

Categorisation of the studies in the review. Table 2 categorises the 23 studies in terms of their study design and the policy effects they identify, and also indicates the sign of the relationship between leave and fertility (final column). Table 2 shows that all 6 studies identifying either the future-child effect or a total effect report positive results, whereas the 17 studies that identify the current-child effect report a mixture of negative, null, and positive results. Table 2 gives a key for each study, to enable quick referencing.

Study quality. We applied the ROBINS-I questionnaire to assess risk of bias (RoB) for all 23 studies. We present our results separately for the current-child effect studies, and for the total effect and future-child effect studies. Before we present our results in Sections "Current-child effect studies" and "Future-child effect and total effect studies", Section "Application of ROBINS-I" explains ROBINS-I in detail, and discusses three considerations that were important for how we adapted ROBINS-I.

Application of ROBINS-I. The original purpose of ROBINS-I is to evaluate RoB in non-randomised studies of medical interventions, so it is not directly applicable to evaluations of the effect of public policies (Sterne et al., 2016). We adapted ROBINS-I in terms of detection bias, assignment to treatment, and in terms of policy "availability" and "eligibility". A detailed discussion of why and how we adapted ROBINS-I is given in Supplementary Appendix D.

Current-child effect studies. We identified four major potential sources of RoB in the current-child effect studies. The first three correspond to ways in which the treatment and control groups may differ systematically, and the fourth concerns the simultaneous introduction of other policies. Details of how we handled these four sources of RoB are discussed in Supplementary Appendix E. Supplementary Appendix E also discusses the results of applying ROBINS-I to the current-child effect studies; these results are summarised in Fig. 4. Two studies (Farré 2007 CC1+ and Lalive 1996 CC1) were found to be at critical RoB, and therefore omitted from the synthesis.

Future-child effect and total effect studies. In the future-child effect and total effect studies, we identified two major potential sources of RoB, in terms of time-varying confounding and long-term fertility trends. The details of these sources of RoB are discussed in Supplementary Appendix F, as is our application of ROBINS-I to these studies. No studies were found to be at critical RoB, and so all studies were included in the synthesis.

Synthesis of results
Current-child effect studies. Almost none of the current-child effect studies found a significant impact of leave on fertility. Out of the 17 current-child effect studies that were evaluated for RoB, two were judged to be at critical RoB overall and are therefore excluded from the synthesis. Two of the 15 remaining studies reported a positive relationship between leave and fertility, and 13 reported no significant relationship. The two studies that found positive results were Dahl 1992 CC1+ and Lalive 1990 CC1, both of which were judged to be at serious overall RoB. Of the 13 studies reporting no significant relationship, three were at low RoB, one was at moderate RoB, and nine were at serious RoB.

Table 2 Categorisation of the studies in this review, in terms of study design and type of effect identified.

Author(s) and year	Study key	Reform year	Study design(s)	Type of effect on fertility	Relationship	Generosity
Ang (2015)	Ang 2006 TO+	2006	Case-control DID	Total effect	Positive	Generous
Cannonier (2014)	Cannonier 1993 FC0	1993	Case-control DID	Future-child effect (parity 0)	Positive	Generous
	Cannonier 1993 T1			Total effect (parity 1)	Positive	Generous
Carneiro, Løken, and Salvanes (2015)	Carneiro 1977 CC1+	1977	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Cools, Fiva, and Kirkebøen (2015)	Cools 1993 CC1+	1993	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Cygan-Rehm (2016)	Cygan-Rehm 2007 CC1+	2007	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Dahl et al. (2016)	Dahl 1987 CC1+	1987	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1988 CC1+	1988		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1989 CC1+	1989		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1990 CC1+	1990		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1991 CC1+	1991		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1992 CC1+	1992		Current-child effect (parity 1+)	Positive	Ungenerous
Duvander, Johansson, and Lappegard (2016)	Duvander 1993 CC1	1993	Short before-after	Current-child effect (parity 1)	Null	Ungenerous
	Duvander 1993 CC2			Current-child effect (parity 2)	Null	Ungenerous
	Duvander 1995 CC1	1995		Current-child effect (parity 1)	Null	Ungenerous
	Duvander 1995 CC2			Current-child effect (parity 2)	Null	Ungenerous
Farré and González (2019)	Farré 2007 CC1+	2007	Short before-after	Current-child effect (parity 1+)	Negative	Ungenerous
Hart, Andersen and Drange (2019)	Hart 2009 CC1+	2009	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Lalive and Zweimüller (2009)	Lalive 1990 CC1	1990	Short before-after	Current-child effect (parity 1)	Positive	Generous
	Lalive 1990 FC1		Long before	Future-child effect (parity 1)	Positive	Generous
	Lalive 1996 CC1	1996	Short before-after	Current-child effect (parity 1)	Null	Ungenerous
	Lalive 1996 FC1		Long before	Future-child effect (parity 1)	Positive	Ungenerous
Raute (2019)	Raute 2007 TO+	2007	Case-control DID	Total effect	Positive	Generous

The "Relationship" column indicates the sign of the relationship between leave and fertility, and a "Null" relationship indicates that no statistically signification relationship was found. The "Generosity" column indicates whether the entitlement changes were large, either in absolute terms or relative to the pre-reform entitlements.

Risk of bias domain

Study reference	1	2	3	4	5	6	7	Overall
Current-child effect studies								
Carneiro 1977 CC1+	+	+	+	+	-	+	+	-
Cools 1993 CC1+	X	+	+	+	-	+	+	X
Cygan-Rehm 2007 CC1+	+	+	+	+	+	+	+	+
Dahl 1987 CC1+	X	+	+	+	-	+	+	X
Dahl 1988 CC1+	X	+	+	+	-	+	+	X
Dahl 1989 CC1+	X	+	+	+	-	+	+	X
Dahl 1990 CC1+	X	+	+	+	-	+	+	X
Dahl 1991 CC1+	X	+	+	+	-	+	+	X
Dahl 1992 CC1+	X	+	+	+	-	+	+	X
Duvander 1993 CC1	X	+	+	+	-	+	+	X
Duvander 1993 CC2	X	+	+	+	-	+	+	X
Duvander 1995 CC1	+	+	+	+	-	+	+	+
Duvander 1995 CC2	+	+	+	+	+	+	+	+
Farré 2007 CC1+	!	+	+	+	X	+	+	!
Hart 2009 CC1+	+	+	+	+	X	+	+	X
Lalive 1990 CC1	X	+	+	+	-	+	+	X
Lalive 1996 CC1	!	+	+	+	-	+	+	!
Future-child effect and total effect studies								
Ang 2006 T0+	X	+	+	+	X	+	+	X
Cannonier 1993 FC0	-	+	+	+	-	+	+	-
Cannonier 1993 T1	-	+	+	+	-	+	+	-
Lalive 1990 FC1	X	+	+	+	-	+	+	X
Lalive 1996 FC1	X	+	+	+	-	+	+	X
Raute 2007 T0+	-	+	+	+	+	+	+	-

Risk of bias level

Symbol	Level
+	Low
-	Moderate
X	Serious
!	Critical

Risk of bias domain

1	Confounding
2	Selection bias
3	Misclassification bias
4	Performance bias
5	Bias from missing data
6	Detection bias
7	Outcome reporting bias

Fig. 4 ROBINS-I evaluations of included studies. Study references are the same as those in the "Study key" column of Table 2. Figure adapted from visualisations produced by the R package "robvis" (McGuinness and Higgins, 2020).

Dahl 1992 CC1+ found that post-reform mothers had 0.042 more children after 14 years, a finding that was significant at the 10% level. In contrast, Lalive 1990 CC1 found that post-reform mothers were 3.5 percentage points likelier to have a birth up to ten years after the reform, a finding that was significant at the 1% level. In this way, the finding of Dahl 1992 CC1+ is very small and possibly only significant due to sampling error, whereas the finding of Lalive 1990 CC1 is much larger and more clearly significant. All of the current-child effect studies are from Northern and Western Europe: 11 are from Norway, two are from Sweden, one is from Germany, and one is from Austria. Case studies from Norway might appear to dominate our findings; however, excluding the Norweigian studies does not affect our conclusions, since three of the four non-Norwegian studies report null results.

Increases in entitlements under a new leave policy can be conceptualised in terms of whether the policy provides a lot more money or length of leave (i.e., absolute generosity), and in terms of whether the entitlement increases are large relative to the pre-

reform entitlements (i.e., relative generosity). The reform of Lalive 1990 CC1 was both absolutely and relatively generous, doubling the length of leave from 12 months to 24 months and remaining at a flat rate of 340 Euros a month. Moreover, the reform entitled women to automatically renew their leave period if they had another birth within 27.5 months of the previous birth (rather than 15.5 months for pre-reform mothers). The renewal entitlement created a strong incentive for post-reform mothers to have another birth quickly relative to pre-reform mothers (a "speed premium"), since it is biologically feasible to have a birth within 27.5 months of a previous birth, but not feasible within 15.5 months. In contrast, the reform of Dahl 1992 CC1+ was neither absolutely generous nor relatively generous. The duration of leave only increased by 3 weeks, from 32 weeks to 35 weeks, and the weekly remuneration rate remained the same.

For 12 of the studies that reported no significant relationship between leave and fertility, increases in leave tended to be between 2–8 weeks with no increase in the rate of remuneration,

so the absolute generosity was low. The relative generosity of the reforms in these studies was also low, since all the pre-reform leave duration was at least 18 weeks. The 13th study with null findings evaluated a moderately generous reform (Carneiro 1977 CC1+), in which maternity leave increased from 12 weeks unpaid to 18 weeks at 100% income replacement, plus 1 year unpaid. None of the null effect studies evaluated reforms that introduced speed premiums as in Lalive 1990 CC1. Overall, it seems that generous new entitlements and speed premiums are preconditions for a leave policy having a current-child effect on fertility. However, there is only one study that meets these preconditions, and so we cannot conclude that generous entitlement increases and speed premiums are sufficient for current-child effect to operate.

Future-child effect and total effect studies. All six studies that evaluate either the future-child effect or the total effect find a positive causal impact of leave on fertility. However, we omit one of these studies from the synthesis—Lalive 1996 FC1—since the authors do not report numerical results for that study.[4] Of the five remaining studies, Ang 2006 T0+, Cannonier 1993 T1 and Raute 2007 T0+ use a case-control DID design to evaluate a total effect; Cannonier 1993 FC0 uses a case-control design to evaluate a future-child effect; and Lalive 1990 FC1 uses a long before design to evaluate a future-child effect. Ang 2006 T0+ and Lalive 1996 FC1 were judged to be at serious RoB, and the other three studies were judged to be at moderate RoB. In terms of effect sizes, Ang 2006 T0+ finds that women benefitting from the reform were 23.53% likelier to have a birth, a finding which is significant at the 1% level. Similarly, Raute 2007 T0+ finds that women benefitting from the reform are 16% likelier to have a birth, a finding which is also significant at the 1% level. Ang 2006 T0+ and Raute 2007 T0+ are the only studies that evaluate the total effect of a leave policy on women at all parities, and therefore provide the best evidence of the impact of leave policies on the aggregate fertility of beneficiaries. The two studies in Cannonier (2014) find the 1993 FMLA reform in the US to make women 5.19 percentage points likelier to have a first birth, and 2.96 percentage points likelier to have a second birth, which are significant at the 5% and 10% level, respectively. Similarly, Lalive 1990 FC1 finds women 6.8 percentage points likelier to have a second birth, which is significant at the 5% level. The effect sizes found in these five studies are all quite large, suggesting that leave policies can potentially have a large impact on increasing fertility. In terms of their geographical coverage, 2 studies are from the US, 1 study is from Canada, 1 from Germany, and 1 from Austria.

All five of the future-child effect or total effect studies evaluate reforms that were either absolutely generous or relatively generous. For both of the two total effect parity 0+ studies, the increase in maximum total benefits over the leave period was roughly $20,000 USD. As discussed in the previous section, the 1990 Austrian parental leave reform in Lalive 1990 FC1 doubled the duration of leave from 12 months to 24 months, and kept the same rate of remuneration. The 1993 FMLA reform evaluated by Cannonier (2014) was not absolutely generous, in that it only granted eligible women 12 weeks of unpaid leave. However, prior to the FMLA women did not have any statutory leave entitlements, and so the FMLA represented a large relative increase. Since all five studies evaluate reforms that provided generous increases in benefits, we cannot establish whether an ungenerous policy would also impact fertility through the future-child or total effect. Generosity may be sufficient but not necessary for a policy to impact fertility; alternatively, generosity may be both sufficient and necessary. However, we can conclude that—at a minimum—generosity is sufficient for a new leave policy to increase fertility.

Discussion

In this section, we discuss four topics: the implications of our findings for pro-natal policy; the implications of our findings for research on gender equity theory and gender roles; the applicability of our conceptual, methodological, and RoB analyses to studies evaluating pro-natal policies; and the limitations of our review.

Our findings suggest that leave policies are a viable strategy for governments seeking to increase fertility in low-fertility settings. Moreover, leave policies can be a cost-effective strategy for increasing fertility, compared with other pro-natal policies. When policymakers discuss the effect of leave on fertility, generally they are referring to either the future-child effect or the total effect—i.e., the extra incentive given to individuals to choose to have a child. The current-child effect is of marginal interest, and it seems that it is overrepresented in the literature since it may be easier to identify empirically. In terms of the academic literature, our analysis and findings might explain why commentators have heretofore evaluated the evidence as "mixed". When treated as an undifferentiated whole, it does seem that the evidence is indeed mixed. However by filtering out studies of the current-child effect from studies of the future-child effect or total effect—and by arguing that the future-child and total effects are the effects important for policymakers wanting to increase aggregate fertility —we have demonstrated that the evidence for the effect of leave on fertility is entirely supportive.

In the Background section, we explained how leave is important for contemporary fertility because of how it alters the gendered pattern of labour, and so here we reflect on whether our findings support gender equity theory. Our motivation for focussing on leave was that gender inequity in the balance of domestic and formal work has been identified as a key driver of low fertility, and that leave can potentially equalise this balance and thereby promote fertility (McDonald, 2000). The future-child effect and total effect studies themselves do not evaluate the extent to which gender equalisation was the mechanism by which those reforms increased fertility; however, other studies on the same reforms may be suggestive. The 2007 parental leave reform in Germany evaluated by Raute (2019) is also studied by Schober (2014), who examines whether the reform was associated with changes in paternal childcare and housework. Using a similar DID design to Raute (2019), Schober (2014) finds that fathers eligible for the new entitlements increased their average weekday childcare hours by 0.6 ($p < 0.05$) after 12 months, and by 0.44 ($p < 0.05$) after 30 months. Moreover, fathers who took leave increased their average weekday childcare hours by 2.45 ($p < 0.05$) after 12 months. However, there was no evidence of a similar increase in paternal housework, limited evidence of a reduction in maternal childcare, and no evidence of a reduction in maternal housework. In light of the reform's impact on the equalisation of childcare, the increase in fertility is consistent with expectations from gender equity theory.

The impact of a leave reform on gender roles is likely to be mediated by the type of reform (maternity, paternity, or parental), and national and temporal context. The reforms evaluated in the future-child effect and total effect studies were implemented at two different times: in 1990–1996 for Austria and the USA, and in 2006–2007 for Canada and Germany. The reforms in the 1990s covered maternity leave, either in law or in practice: the parental leave in the Austrian reforms was overwhelmingly taken by mothers, and with no time reserved for fathers (Lalive and Zweimüller, 2009). By contrast, both the German and Canadian reforms were parental, and included father's quotas (Ang, 2015; Raute, 2019). The Austrian and American reforms therefore may have acted to reinforce the male-breadwinner model, whereas the German and Canadian reforms may have promoted a gender-

equal model (Farré, 2016; Barnes, 2015; Blome, 2016). However, it is unclear whether the gendered pattern of leave reforms encourage gendered behaviour, or whether both reforms and behavioural changes are reflective of changes in gender ideology. Kang (2019) notes that OECD countries experienced a convergence in the gender equality of family policies from 1990–2010, with the social democratic welfare states of Norway, Sweden and Denmark providing the most gender-egalitarian family policies by 2010.

The conceptual, methodological, and RoB sections of this review provide a clear and logical framework for conducting, classifying, and evaluating studies of the effect of leave on fertility. The bifurcation of studies as identifying either the current-child or future-child effect separates studies into two fundamentally different types, meaning that studies evaluating different types of effects cannot be directly compared (Lalive and Zweimüller, 2009). Furthermore, our framework classifies studies in terms of the parity of the individuals they analyse. For future empirical research, this framework will enable researchers to better conceptualise and understand the precise effects identified by their studies. This conceptual framework has been developed in order to categorise studies of the effect of leave on fertility, but it could be used to analyse the effects of other pro-natal policies on fertility. Moreover, the study designs associated with evaluating these effects—the short before-after design, the long before design, and the case-control DID design —could also be used for to evaluate the effects of other pro-natal policies.

In terms of RoB, we hope that our analysis in section "Study quality" could provide the foundations for a custom-built tool for assessing RoB in studies of public policy evaluations. Historically, systematic review methods in social science have been adapted from pre-existing methods in medical research, and we hope that our analysis will be used in this tradition. Our application of ROBINS-I also shows that case-control DID studies of eligibility are generally at lower RoB than long before studies or studies of leave availability, suggesting that researchers should choose to use case-control DID evaluations of eligibility where possible.

There are three considerations that may limit the applicability of our findings to other settings. Firstly, the number of studies of the future-child effect or the total effect are quite small, meaning that it is difficult to establish the relative importance of leave generosity. Specifically, all five studies evaluate generous reforms, and so we cannot know whether ungenerous entitlement increases would affect fertility. Secondly, two of the future-child and total effect studies evaluate reforms that were implemented in the recent past (2006 and 2007), and so whether these reforms will have an impact on completed fertility remains to be seen. Lastly, the geographical coverage of these studies is limited to Northern and Western Europe, and North America. This means that generalising these findings to other low-fertility settings (such as Southern Europe, Eastern Europe, and East Asia) may not be appropriate.

Conclusion

In this review, we sought to examine what the best available evidence showed about the effect of leave on fertility. Our motivation was that fertility is very low in many countries and declining in many more, that national governments devote large resources to increase fertility through family policies, and that most academic commentators argue that the evidence for the effect of leave on fertility is mixed. Our focus on leave was motivated by the way maternity, paternity and parental leave change the gender distribution of formal and domestic labour, a factor identified as important in causing low fertility. In conducting the review, we followed a review protocol written prior to searching the literature, focussed only on primary empirical

studies with experimental and quasi-experimental designs, and evaluated the quality of studies using the ROBINS-I framework.

We identified 23 studies, which examined the impact of leave policies on fertility. In order to understand the seemingly contradictory findings of these studies, we extended Lalive and Zweimüller's (2009) conceptual framework. Our extension to this framework enabled us to categorise the studies based on the effects identified, the parity of study participants, and the study designs used to identify effects. This categorisation demonstrated that all of the studies with null or negative findings were only identifying a narrow type of effect of leave on fertility (the current-child effect), an effect, which accounts for only a small part of the total effect. Moreover, we argued that the current-child effect is likely to be of marginal interest to policymakers considering leave as a means of increasing fertility. Our categorisation also demonstrated that studies identifying a more complete effect (either the future-child effect or the total effect) all had positive and significant findings. Moreover, the effect sizes found in these studies were large, with the probability of a next birth increased by as much as 24%. We therefore reject the contention that the evidence for the effect of leave on fertility is mixed. Rather, we find that the apparently mixed evidence is simply an artefact of sub-optimal study design. The results contribute to the understanding of the effect of leave on fertility by showing that different study designs can only identify certain types of effects, and by showing leave can significantly increase fertility when increases in benefits are generous. Using supplementary evidence on the impact of the same policies on increasing paternal childcare, it seems plausible that leave may increase fertility partly through equalising domestic gender roles, as suggested by gender equity theory.

Notes

1 In some countries such as Sweden, gender equality is an explicit goal of leave policy (Duvander and Johansson 2012).
2 By "different studies" we mean individual models (or sets of related models) that are different either in terms of the specific policy being evaluated, or different in terms of the analytical sample (e.g., a paper that has separate models for the effect of a policy on women at parity 1 and parity 2 counts as having two different studies).
3 Alternatively, Lalive and Zweimüller's total effect (for women at parities of one or more) can be regarded as an underestimate, or lower bound, on the total effect (for women at parities of zero or more).
4 The only mention of the results of Lalive 1996 FC1 in Lalive and Zweimüller (2009) is in footnote 32, page 1399.

References

Ang XL (2015) The effects of cash transfer fertility incentives and parental leave benefits on fertility and labor supply: evidence from two natural experiments. J Fam Econ Issues 36(2):263–288

Angrist JD, Pischke J-S (2009) Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Princeton

Baizan P, Arpino B, Eric Delclos C (2016) The effect of gender policies on fertility: the moderating role of education and normative context. Eur J Popul 32:1–30

Balbo N, Billari FC, Mills M (2013) Fertility in advanced societies: a review of research. Eur J Popul 29:1–38

Barnes MW (2015) Gender differentiation in paid and unpaid work during the transition to parenthood. Sociol Compass 9(5):348–364. https://doi.org/10.1111/soc4.12263

Baum CL, Ruhm CJ (2016) The effects of paid family leave in California on labor market outcomes. J Policy Anal Manag 35(2):333–U135

Baxter J, Hewitt B, Haynes M (2008) Life course transitions and housework: marriage, parenthood, and time on housework. J Marriage Family 70(May):259–272

Beaujouan E, Berghammer C (2019) The gap between lifetime fertility intentions and completed fertility in Europe and the united states: a cohort approach. Popul Res Policy Rev 38:507–535

Becker GS (1973) A theory of marriage: part I. J Polit Econ 81(4):813–846

Bergsvik J, Fauske A, Hart RK (2020) Effects of policy on fertility: a systematic review of (quasi)experiments. Statistics Norway, Research Department, Oslo (Discussion Papers No. 922, Feb 2020)

Blome A (2016) Normative beliefs, party competition, and work-family policy reforms in Germany and Italy. Comp Polit 48(4):479–503. https://doi.org/10.5129/001041516819197610

Bloom DE, Canning D, Fink G, Finlay JE (2010) The cost of low fertility in Europe. Eur J Popul 26(2):141–158

Budig MJ, Misra J, Boeckmann I (2015) Work–family policy trade-offs for mothers? Unpacking the cross-national variation in motherhood earnings penalties. Work Occup 43(2):119–177. https://doi.org/10.1177/0730888415615385

Bünning M (2015) What happens after the 'Daddy Months'? fathers' involvement in paid work, childcare, and housework after taking parental leave in germany. Eur Sociol Rev 31:738–748. https://doi.org/10.1093/esr/jcv072

Buttner T, Lutz W (1990) Estimating fertility responses to policy measures in the German Democratic Republic. Popul Dev Rev 16(3):539–555

Caldwell JC, Caldwell P, McDonald P (2002) Policy responses to low fertility and its consequences: a global survey. J Popul Res 19(1):1–24

Campbell Collaboration (2019) Campbell systematic reviews: policies and guidelines. Campbell Policies and Guidelines Series No. 1. https://doi.org/10.4073/cpg.2016.1

Cannonier C (2014) Does the Family and Medical Leave Act (FMLA) increase fertility behavior? J Labor Res 35:105–132

Casterline JB, Han S (2017) Unrealized fertility: Fertility desires at the end of the reproductive career. Demogr Res 36:427–453

Chen M, Yip PSF (2017) The discrepancy between ideal and actual parity in Hong Kong: fertility desire, intention, and behavior. Popul Res Policy Rev 36(4):583–605

Ciccia R, Verloo M (2012) Parental leave regulations and the persistence of the male breadwinner model: Using fuzzy-set ideal type analysis to assess gender equality in an enlarged Europe. J Eur Soc Policy 22(5):507–528. https://doi.org/10.1177/0958928712456576

Cools S, Fiva JH, Kirkebøen LJ (2015) Causal effects of paternity leave on children and parents. The. Scand J Econ 117(3):801–828

Cygan-Rehm K (2016) Parental leave benefit and differential fertility responses: evidence from a German reform. J Popul Econ 29(1):73–103

Dahl GB, Løken KV, Mogstad M, Vea Salvanes K (2016) What is the case for paid maternity leave? Rev Econ Stat 98(4):655–670

Dearing H (2015) Does parental leave influence the gender division of labour? Recent empirical findings from Europe. Institut für Sozialpolitik, WU Vienna University of Economics and Business. Vienna. Retrieved from https://epub.wu.ac.at/id/eprint/4505

Duvander A-Z, Johansson M, Lappegard T (2016) Family Policy Reform Impact on Continued Fertility in the Nordic Countries. Stockholm University Linnaeus Center on Social Policy and Family Dynamics in Europe, Stockholm (Working Paper 2016: 2)

Duvander A-Z, Lappegård T, Andersen SN, Garðarsdóttir Ó, Neyer G, Viklund I (2019) Parental leave policies and continued childbearing in Iceland, Norway, and Sweden. Demogr Res 40(51):1501–1528

Duvander AZ, Johansson M (2012) What are the effects of reforms promoting fathers' parental leave use? J Eur Soc Policy 22(3):319–330

Esping-Andersen G (1990) The Three Worlds of Welfare Capitalism. Princeton University Press, Princeton

Ermisch JF (2003) An economic analysis of the family. Princeton University Press, Princeton

Evertsson M, Duvander A-Z (2011) Parental leave–possibility or trap? does family leave length effect Swedish Women's Labour Market Opportunities? Eur Soc Rev 27(4):435–450. https://doi.org/10.1093/esr/jcq018

Eydal GB, Rostgaard T (2018) Introduction to the Handbook on Family Policy. In: Eydal GB, Rostgaard T (eds) Handbook on family policy. Edward Elgar, Cheltenham, UK, pp. 2–9

Farré L (2016) Parental leave policies and gender equality: a survey of the literature. Estud Econ Apl 34:45–60

Farré L, González L (2019) Does paternity leave reduce fertility? J Public Econ 172:52–66

Fernández-Cornejo JA, Pozo-García ED, Esco L, Castellanos-Serrano C (2018) Can an egalitarian reform in the parental leave system reduce the motherhood labor penalty? Some evidence from Spain. Rev Esp Sociol 27:45–64. https://doi.org/10.22325/fes/res.2018.33

Gauthier AH (2007) The impact of family policies on fertility in industrialized countries: a review of the literature. Popul Res Policy Rev 26:323–346

Gauthier AH (2008) Some theoretical and methodological comments on the impact of policies on fertility. Vienna Yearb Popul Res 6:25–28. https://doi.org/10.1553/populationyearbook2008s25

Gauthier AH, Bartova A (2018) The impact of leave policies on employment, fertility, gender equality, and health. In: Shockley KM, Shen W, Johnson RC (eds) The Cambridge handbook of the global work family interface. Cambridge University Press, Cambridge, pp. 120–138

Gauthier AH, Koops JC (2018) The history of family policy research. In: Eydal GB, Rostgaard T (eds) Handbook of family policy. Edward Elgar, Cheltenham, pp. 11–23

Gauthier AH, Philipov D (2018) Can policies enhance fertility in Europe? Vienna Yearb Popul Res 6:1–16. https://doi.org/10.1553/populationyearbook2008s1

Geyer J, Haan P, Wrohlich K (2015) The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. Labour Econ 36:84–98. https://doi.org/10.1016/j.labeco.2015.07.001

Goldscheider F, Bernhardt E, Lappegård T (2015) The gender revolution: a framework for understanding changing family and demographic behavior. Popul Dev Rev 41(2):207–239

Haas L (2003) Parental leave and gender equality: lessons from the European Union. Rev Policy Res 20(1):89–114. https://doi.org/10.1111/1541-1338.d01-6

Hart RK, Andersen SN, Drange N (2019) Effects of extended paternity leave on union stability and fertility. Statistics Norway Discussion Papers, 899

Hegewisch A, Gornick JC (2011) The impact of work-family policies on women's employment: a review of research from OECD countries. Commun Work Fam 14:119–138. https://doi.org/10.1080/13668803.2011.571395

Higgins JPT, Green S (eds.) (2011) Cochrane handbook for systematic reviews of interventions (5.1.0). Retrieved from www.handbook.cochrane.org

Hoem JM (1990) Social policy and recent fertility change in Sweden. Popul Dev Rev 16(4):735–748

Hoem JM (1993) Public policy as the fuel of fertility: effects of a policy reform on the pace of childbearing in Sweden in the 1980s. Acta Sociol 36(1):19–31

Hoem JM (2005) Why does Sweden have such high fertility? Demogr Res 13:559–572

Hoem JM (2008) The impact of public policies on European fertility. Demogr Res 19(10):249–260

Hong SH, Sullivan R (2016) The effects of subsidies for childbearing on migration and fertility: evidence from Korea. Singapore Econ Rev 61(4):1–16

Kang JY (2019) Convergence of family policy across welfare regimes (1990 to 2010): Different connotations of family policy expansion. Int J Soc Welfare 28(2):167–178. https://doi.org/10.1111/ijsw.12331

Keng S-H, Sheu S-J (2011) Can education subsidy reverse falling fertility? The Case of Taiwan. J Deve Stud 47(11):1677–1698. https://doi.org/10.1080/00220388.2011.609585

Kotsadam A, Finseraas H (2011) The state intervenes in the battle of the sexes: Causal effects of paternity leave. Soc Sci Res 40(6):1611–1622

Lalive R, Zweimüller J (2009) How does parental leave affect fertility and return to work? Evidence from two natural experiments. Q J Econ 124(3):1363–1402

Lappegård T (2010) Family policies and fertility in Norway. Eur J Popul 26(1):99–116

Lechner M (2010) The estimation of causal effects by difference-in-difference methods. Found Trends Econom 4(3):165–224. https://doi.org/10.1561/0800000014

Matysiak A, Szalma I (2014) Effects of parental leave policies on second birth risks and women's employment entry. Population 69(4):599–636

McDonald P (2000) Gender equity in theories of fertility transition. Popul Dev Rev 26(3):427–439. https://doi.org/10.1111/j.1728-4457.2000.00427.x

McDonald P (2006) Low fertility and the state: the efficacy of policy. Popul Dev Rev 32(3):485–510

McDonald P (2008) Very low fertility: consequences, causes and policy approaches. Japan J Popul 6(1):19–23

McGuinness LA, Higgins JPT (2020) Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. Res Synth Method 2020:1–7

Meier V, Rainer H (2017) Daddy months. J Popul Econ 30:875–892

Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339:b2535

Monnier A (1990) The effects of family policies in the German Democratic Republic: a re-evaluation. Population (English Selection) 2:127–140

Neilson J, Stanfors M (2014) It's about time! Gender, parenthood, and household divisions of labor under different welfare regimes. J Family Issue 35:1066–1088. https://doi.org/10.1177/0192513X14522240

Neyer G, Andersson G (2008) Consequences of family policies on childbearing behavior: effects or artifacts? Popul Dev Rev 34(4):699–724

Nieuwenhuis R, Lancker WV (2020) Introduction: a multilevel perspective on family policy. In: Nieuwenhuis R, Lancker WV (eds) The Palgrave handbook of family policy. Palgrave Macmillan, Cham, Switzerland, pp. 3–24

OECD (2019a) Family benefits public spending (indicator). https://doi.org/10.1787/8e8b3273-en

OECD (2019b) Gross domestic product (GDP) (indicator). https://doi.org/10.1787/dc2f7aec-en

OECD (2019c) PF1.1: Public spending on family benefits. OECD Family Database. http://www.oecd.org/els/family/database.htm

OECD (2019d) PF2.1: Parental leave systems. OECD Family Database. http://www.oecd.org/els/family/database.htm

Olivetti C, Petrongolo B (2017) The economic consequences of family policies: lessons from a century of legislation in high-income countries. J Econ Perspect 31(1):205–230

Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M et al. (2006) Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC methods programme. Lancaster University

Priebe J (2020) Quasi-experimental evidence for the causal link between fertility and subjective well-being. J Popul Econ 33(3):839–882

Pronzato CD (2009) Return to work after childbirth: does parental leave matter in Europe? Rev Econ Househ 7(4):341–360

Raute A (2019) Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits. J Public Econ 169:203–222

Raybould A, Sear R (2020) Children of the (gender) revolution: A theoretical and empirical synthesis of how gendered division of labour influences fertility. Popul Stud 75(2):169–190. https://doi.org/10.1080/00324728.2020.1851748

Rindfuss RR, Choe MK (eds.) (2016) Low fertility, institutions, and their policies: variations across industrialized countries. Springer, Switzerland

Salles A (2006) The effects of family policy in the former GDR on nuptiality and births outside marriage. Population 61(1-2):141–151

Saxonberg S (2013) From defamilialization to degenderization: toward a new welfare typology. Soc Policy Adm 47(1):26–49. https://doi.org/10.1111/j.1467-9515.2012.00836.x

Schober PS (2014) Parental leave and domestic work of mothers and fathers: a longitudinal study of two reforms in West Germany. J Soc Policy 43:351–372. https://doi.org/10.1017/S0047279413000809

Shadish WR, Cook TD, Campbell DT (2002) Experimental and quasi-experimental designs for generalized causal inference (3rd edn). Houghton Mifflin Company, Boston

Spéder Z, Kapitány B (2014) Failure to realize fertility intentions: a key aspect of the post-communist fertility transition. Popul Res Policy Rev 33(3):393–418

Spiess CK, Wrohlich K (2008) The parental leave benefit reform in Germany: costs and labour market outcomes of moving towards the Nordic Model. Popul Res Policy Rev 27:575–591. https://doi.org/10.1007/s11113-008-9086-5

Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M et al. (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ (2016) 355:i4919

Stropnik N, Sircelj M (2008) Slovenia: generous family policy without evidence of any fertility impact. Demogr Res 19(26):1019–1058

Tamm M (2019) Fathers' parental leave-taking, childcare involvement and labor market participation. Labour Econ 59:184–197

Thévenon O (2011) Family policies in OECD countries: a comparative analysis. Popul Dev Rev 37(1):57–87. https://doi.org/10.1111/j.1728-4457.2011.00390.x

Thomas J, Rowe F, Lin E (2019) The effect of parental leave policies on increasing fertility: a systematic review. In: PROSPERO International prospective register of systematic reviews. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019128493. Accessed 19 Jul 2022

Ugur ZB (2020) Does having children bring life satisfaction in Europe? J Happiness Stud 21(4):1385–1406

UN DESA (2019) World population prospects 2019: Highlights. Department of Economic and Social Affairs, World Population Prospects 2019, United Nations

Wanqing Z (2021) China Mulls longer parental leave to encourage more births. https://www.sixthtone.com/news/1007712/china-mulls-longer-parental-leave-to-encourage-more-births. Accessed 19 Aug 2021

Zabel C (2009) Eligibility for maternity leave and first birth timing in Great Britain. Popul Res Policy Rev 28:251–270. https://doi.org/10.1007/s11113-008-9098-1

Acknowledgements

Author contributions

JT: made substantial contributions to the conception and design of the study; drafted the study and revised it critically for important intellectual content; gave final approval of the version to be published; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated or resolved. FR: made substantial contributions to the conception and design of the study; drafted the study and revised it critically for important intellectual content; gave final approval of the version to be published; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated or resolved. PW: made substantial contributions to the conception and design of the study; drafted the study and revised it critically for important intellectual content; gave final approval of the version to be published; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated or resolved. ESL: made substantial contributions to the conception and design of the study; drafted the study and revised it critically for important intellectual content; gave final approval of the version to be published; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated or resolved.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-022-01270-w.

Correspondence and requests for materials should be addressed to Jac Thomas.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.