Can Policies Stall the Fertility Fall? A Systematic Review of the (Quasi-) Experimental Literature

Janna Bergsvik, Agnes Fauske and Rannveig Kaldager Hart

In the course of the twentieth century, social scientists and policy analysts have produced a large volume of literature on whether policies boost fertility. This paper describes the results of a systematic review of the literature on the effects of policy on fertility since 1970 in Europe, the United States, Canada, and Australia. Empirical studies were selected through extensive systematic searches, including studies using an experimental or quasi-experimental design. Thirty-five studies were included, covering reforms of parental leave, childcare, health services, and universal child transfers. In line with previous reviews, we find that childcare expansions increase completed fertility, while increased cash transfers have temporary effects. New evidence on parental leave expansions, particularly from Central Europe, suggests larger effects than previously established. High-earning couples benefit more from parental leave expansions, while expanding childcare programs can reduce social inequalities on other domains. Subsidizing assisted reproductive treatments shows some promise of increasing birth rates for women over the age of 35. Countries that to date have limited support for families can build on solid evidence if they choose to expand these programs.

Introduction

The decline of fertility below replacement levels has been met with concern in several advanced economies, not least because it contributes to an aging workforce and challenges in caring and providing for tomorrow's

Janna Bergsvik is Researcher, Research Department, Statistics Norway, Oslo, Norway. Agnes Fauske is Doctoral Research Fellow, Department of Sociology and Human Geography, University of Oslo, Oslo, Norway. Rannveig Kaldager Hart is Researcher, Centre for Evaluation of Public Health Measures and Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway.
E-mail: rannveigkaldager.hart@fhi.no

elderly (McDonald 2006; Neyer et al. 2017; Sobotka et al. 2020). Following the 2007 economic recession, fertility levels have fallen further throughout Europe (Vignoli et al. 2020), with recent studies suggesting shrinking family sizes rather than just postponement effects (Hellstrand et al. 2020). In 2017, the fertility in 83 of 201 countries was below the replacement level (United Nations 2018). Many of these countries also allocate large budget shares to different forms of family support. In particular, 66 percent of European and almost 40 percent of Asian governments had in place policies with the explicit aim of raising fertility or at least impeding a further decline (United Nations 2018).

Countries with extensive public support to families also tend to have higher fertility (e.g., Gauthier and Hatzius 1997; Kalwij 2010; Wood et al. 2016; Gauthier 2007), and a country's fertility often goes up when family policies become more generous (Hoem 1990; Rønsen and Skrede 2010; Buttner and Lutz 1990; McDonald 2006). These observations, in combination with a theoretical understanding that people have more children when costs of childrearing are lower (Becker 1991; Hotz et al. 1997), have led to influential theories linking policy and fertility. One particularly widespread notion is that policies helping families balance two jobs and childrearing, and thus reducing the opportunity costs, contribute to high fertility. Increasing father involvement may further reduce the mother's opportunity costs, potentially increasing (her) fertility preferences (Goldscheider et al. 2015).

The efficacy of public policies to influence fertility has nevertheless been debated, not least because of challenges to disentangling the causal effects of policies from other characteristics of individuals and societies, such as values, the degree of family-friendliness, and economic conditions (Hoem 2008; Neyer and Andersson 2008). Gauthier concludes in her extensive review on policy effects that "knowledge on this matter is still limited, and calls for complex modeling of the causal relationship" (2007, 142). Similarly, McDonald concludes that policy effects will "be difficult to disentangle (…) when there is no counterfactual: what would have happened without the policy change?" (2006, 501–502). The "credibility revolution" in microeconomics was a leap forward in empirical counterfactual analysis (Angrist and Pischke 2010). Fine-grained comparisons across space and time or different population groups are used as naturally occurring "experiments," yielding clearly defined control and treatment groups.

Despite their relevance for this core debate in demography, no systematic overview of studies of causal effects of policies on fertility exists to date. This paper fills this gap by reviewing and synthesizing experimental and quasi-experimental studies of fertility effects of policies. Starting from a broad systematic search, we summarize effects of a wide range of policies with potential relevance for fertility, including childcare, parental leave, health services, and cash transfers. Geographically, our review includes studies from welfare regimes in Europe and North America; the

Central European, Nordic, Anglo-Saxon, and Southern European regimes (see e.g., Esping-Andersen 1990; Sobotka et al. 2020). They all share the experience of an increasing female labor supply and falling fertility over the last half century. The Nordic and to some extent the Central European contexts have seen fertility recuperations, arguably linked to expansion of policies that support dual-earner families. Although the Anglo-Saxon regime is distinct with until recently high, yet socially polarized, fertility and low public support, evidence from policies in this context offers important lessons about the nature of policy effects. We pay careful attention to effect variation by welfare regime, as a background for a systematic discussion of how reform effects depend on the larger societal structure.

We also discuss potential sources of bias in the studies included in our review, including an examination of our material for publication bias. We contrast findings from this literature review with conclusions from earlier reviews of the general fertility-policy literature. Our findings underpin the fact that childcare expansions have lasting impact on fertility, whereas the impact of transfers is substantial yet transitory. By leaning on updated evidence and giving most weight to large reforms evaluated in credible analytical designs, we conclude—in contrast to previous reviews—that large parental leave expansions also have substantial and lasting fertility effects.

Theoretical framework

Attempts to politically influence norms and preferences regarding fertility choice are typically seen as both difficult and, in liberal democracies, largely unacceptable (Schultz 2015). Policies meant to affect fertility will instead be aimed at the economic and time costs of parenting. This understanding of fertility choice fits well with the microeconomic theory of the family, much used as a framework for quasi-experimental analysis. Here, the number of children couples choose to have depends on their purchasing power, the time and money (including foregone income) expected to be needed for childrearing, and parents' preferences for spending that time and money on children rather than using it for other purposes (Becker 1991). Policies affect parents' time and budget constraints. For instance, reducing the price of health care or housing for families with children will lower the cost of raising a child. Increased income (through transfers or tax breaks) will translate into larger family sizes unless parents instead spend more resources on each child. Moreover, policies that increase wages and/or strengthen attachment to the labor market increase the losses arising from taking time off work to care for children (the opportunity cost). The total effect of such policies on fertility could thus be positive or negative.

Policies may also influence fertility by redistributing the time cost of childbearing between the parents. This cost has been incurred disproportionally by mothers, and policies aiming to shift some of the cost to fathers

may strengthen mothers' interest in having a(nother) child (e.g., Goldscheider et al. 2015). The total impact of this is not clear, however, as fathers' fertility desires may be lower when their opportunity costs increase (Farré and González 2019).

One would expect substantial within-population variation in policy responses across social groups (Hakim 2003). A simple example is that reduced kindergarten fees reduce expected childcare costs, but not for families with a preference for parental care over formal care. Another example concerns information. Because individuals have incomplete information about the costs and benefits of (further) childbearing, they have to rely on their perception and expectations of the costs (Goldthorpe 1998). Patterns of information-seeking may vary by socioeconomic status, or between parents and childless adults. Importantly, fertility is also influenced by norms, fecundity, and costs of contraception and abortion (Crimmins 1985). When the last two fall, one would expect that fertility also declines.

Scope and search strategies

Policy scope

Our starting point is that fertility is a private choice that is influenced by the context provided by public policies (Schultz 2015). We do not consider the literature on abortion rights or availability of contraception specifically, but consider the effects on fertility when the cost of contraception is changed as part of a package of changing health service costs. Beyond that, we include relevant policies regardless of their aim, be it fertility increase or simply cutbacks driven by budget deficits. The policies found in the included literature encompass childcare, parental leave, child transfers, and health services.[1] Our search also yielded a substantial number of evaluations of welfare policies, conditional transfers in liberal regimes, and cutbacks in these. Cutbacks, in particular, are often aimed explicitly at reducing (nonmarital) fertility and fertility among welfare dependents. It is difficult to envision how these policies could be "reversed" to become tools to increase fertility; it seems unlikely that a country would implement a package aiming to increase fertility while simultaneously increasing welfare dependency and reducing female labor supply. For completeness, the welfare reforms that satisfied our other selection criteria are described in Appendix A in the Supporting Information. Beyond that, these reforms are not included in our quantitative or narrative synthesis.

Methodological scope

Individual-level studies of policy *uptake* and subsequent fertility behavior could provide misleading estimates of policy effects. For example, a positive

relationship between the uptake of paternity leave and subsequent fertility may reflect an effect of parental leave, but also the fact that family-oriented men have more children and take longer leave with each child (Neyer and Andersson 2008). Quasi-experimental designs handle such self-selection by comparing (conditionally) identical groups exposed to different policies (Angrist and Pischke 2008). Such groups can emerge naturally if a policy is applied to one region but not another (see Neyer and Andersson 2008 for demographic examples), or if benefit levels increase for those with a high income but not for those with a lower income.

Methods for causal inference have developed in parallel in statistics (Holland 1986), econometrics (Angrist and Pischke 2010), and epidemiology (Greenland 1990) but have become particularly dominant within empirical microeconomics. As of 2017, quasi experiments were used in 47 percent of the preprints available on the National Bureau of Economic Research server (Ruhm 2019). In Box 1, we include a brief description of the designs of the studies included in our review, including the identification assumptions they hinge on. These designs are: randomized experiments, regression discontinuity (RD) designs, difference-in-differences (DiD) estimation, two-way fixed-effects (2WFE) panel regression models, and instrumental variable (IV) designs.[2]

Box 1
(Quasi-)experimental methods used in the studies included in the review.

Randomized experiments where a benefit is randomly given to some persons (treatment group) and not to others (control group), provide the most obvious opportunity for evaluating the causal effect of that benefit. However, for practical and ethical reasons experiments are rare, and external validity may be limited if experiments create superficial settings.

Regression discontinuity (*RD*) *designs* are suitable when arbitrary cutoffs (rather than individual choice) define who is affected by a policy, for instance when parents of children born after a given date get longer parental leave. In such a case, those being just ineligible should be similar to those being just eligible and therefore constitute a good comparison group. Visual inspection of the discontinuity in plots is an important test for internal validity.

Difference-in-differences (*DiD*) *estimation* builds on the same logic. Some groups or units are exposed to policy changes or expansions, while others are not. Within-group fertility changes over time are then compared across the groups. Studies should document that pre-trends are parallel in the intervention and control group prior to the intervention. Sometimes a subgroup less impacted by the reform can be identified within the treatment group. For instance, very high-earning women are likely

unmoved by eligibility for a stay-at-home subsidy. If such a third axis exists, a "triple-differenced design (DiDiD)" can be estimated.

Two-way fixed-effects (*2WFE*) *panel regression models* are a generalized form of difference-in-differences (DiD) estimation. By using time and unit fixed effects these models effectively control for confounding shared time shocks and time-constant differences between units. If there are changes over time, other than the reform, that differ between the treatment and control areas, results will be biased.

Instrumental variable (*IV*) *designs* are estimated in two steps using reforms as instruments to generate exogenous (random) variation in an otherwise endogenous regressor. This allows for quantification of reform effects, for instance giving the change in fertility per $100 in additional transfers. Such quantification facilitates comparison across different studies and reforms.

Some characteristics of policy reforms and quasi-experimental designs could, however, give a downward bias of effect sizes. If policy changes happen in small increments, the effect of each increment may be so slight that even full population data have insufficient power to detect it. This holds even if the total effect of the policy is substantial. Further, quasi-experimental designs measure the effect on the individuals who change their behavior because of a reform, for instance fathers who take more leave if and only if the paternity quota is extended. If only a small subgroup is affected by a reform, large reform effects on fertility are unlikely. Last, reforms can change the cost of future children and/or the cost of children already born ("current child effect"). Quite often, reforms change the cost of future children, but the analytical design compares the fertility responses between parents that receive different benefits for current children only. In such cases, the estimated reform effect is likely substantially smaller than the total reform effect (see Lalive and Zweimüller 2009; Raute 2019).

Literature search and selection

Our systematic review comprises studies that used fertility as the outcome, a policy as the key explanatory variable, and (quasi-)experimental methods. The search was further restricted according to several pre-set criteria about the study population, type of policy, comparison groups, outcome, and study design (see Table 1). Based on an understanding that the relationship between policy change and fertility choice happens at the microlevel (see Neyer and Andersson 2008), we included only studies using microdata. To avoid any bias toward the near and known, we conducted keyword searches in several databases and literature list screenings. These searches generated 17,228 unique hits, of which titles and abstracts were screened for relevance in line with the criteria given in Table 1 for literature in

TABLE 1 PICOS for inclusion and exclusion

Criteria	Inclusion	Exclusion
Participants (population)	(1) Populations of nations fully located in Europe (excluding, e.g., Turkey and Russia), North America (Canada and the United States), and Australia. (2) Women or men of childbearing age during the intervention.	(1) Teenage pregnancies. (2) Romania, because of a particularly coercive pro-natalist regime under Ceauşescu that may generally limit external validity.
Interventions	(1) Intervention is a policy implemented at national, regional, or local level. (2) Intervention happened after 1970. (3) The intervention affects the fertility choices of the population.	(1) The intervention directly limits participants' free choice by restricting access to contraception or abortion. (2) The effects of the intervention on fertility are unduly complex or indirect, making the intervention an obviously ineffective means of influencing fertility.
Comparator/ control	(1) The introduction/ revocation of a policy is compared to the absence/ presence of the same policy. (2) Modifications of a policy are compared to the same policy in its previous form. (3) Two different policy treatments are compared.	
Outcomes	(1) Birth rates measured at aggregate (subnational) level. (2) Birth probabilities measured at individual level. (3) Period ("timing") measures. (4) Cohort ("quantum") measures.	(1) Outcome is measured at the country level.
Study design	(1) Field experiments. (2) Quasi-experiments: difference-in-differences, regression discontinuity and instrumental variable design, and any combination of these. (3) Two-way fixed effects, or area fixed effects with detailed controls for period and cohort.	(1) Observational studies that do not use the strategies mentioned for causal identification. (2) Fixed effects are measured at a higher level than treatment.

NOTE: For further details, see protocol (Fauske et al. 2020).

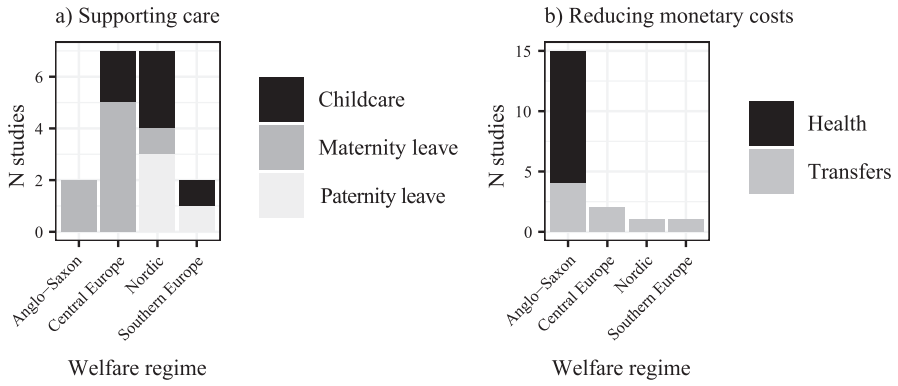
English, French, Spanish, German, Italian, Swedish, Norwegian, and Danish. In total, two researchers read 335 articles in full text to evaluate whether the study design and method were within the scope of the review. Thirty-five articles constituted the final sample. The screening process is illustrated in Figure A1 of Appendix A in the Supporting Information. Details can be found in the protocol (Fauske et al. 2020). The project is also preregistered at PROSPERO (Hart et al. 2019a).

The reforms in context

Effects of the policies investigated may be influenced by the broader setting and policy constellation in which they are implemented, including the degree of economic inequality (Thomson et al. 2014). Importantly, the counterfactual of a reform effect will differ substantially by study context. The effect of one additional year of parental leave will likely depend on whether the alternative is public childcare or (unpaid) care at home. Beyond that, the societal context also influences what kind of policies are implemented in the first place (Neyer and Andersson 2008). In line with previous reviews (Gauthier 2007; Sleebos 2003; Sobotka et al. 2020), we cluster the study countries in welfare regimes to describe the wider institutional and cultural circumstances in which the reforms are carried out (see Esping-Andersen 1990 for an early formulation of the typology).

Of the studies included in our review, the broadest evidence on family policies comes from the Nordic and Central European regimes. Evidence is scarce from Southern Europe. Studies from the Anglo-Saxon countries are relatively many but narrow in focus, concentrated on specific polices (health) or regions (Quebec). No studies from Eastern Europe and Australia fit the inclusion criteria. Figure 1 shows the distribution of studies by welfare regime and topic.

FIGURE 1 Studies by reform and context



The Nordic context

The social–democratic regime is characterized by universal state support for children and (working) parents, and a high commitment to gender equality. From the Nordic context, we draw evidence on childcare, cash transfers, and parental leave (Figure 1). The Nordic welfare states aim to support lifelong, full-time work for men and women alike, and the combination of work and children is today enabled by long paid parental leaves of about one year (Dahl et al. 2016) and widely available public childcare for children above that age (Rindfuss et al. 2010). Fathers' quotas to parental leave aim to ensure that the combination of caregiving and paid work is shared more equally between the parents (Cools et al. 2015).

The Central European context

The conservative regime is characterized by medium levels of financial support for families that tend to relate to a more traditional division of labor, with a subsidiary principle where the state only interferes when the family does not manage on its own (Esping-Andersen 1990). This cluster includes Central European countries such as Germany, France, Belgium, and Austria (Gauthier 2002). Over time, these countries have gradually shifted toward supporting mothers' employment through implementing polices inspired by those in the Nordic countries in a context of different gender roles and family arrangements. Belgium and France were early to provide public childcare (Sobotka et al. 2020). The German unification merged the West German male breadwinner model with the East German dual-earner model (Leitner et al. 2008) and since 2007 substantial improvements in public childcare and changes in parental leave support the dual-earner model in both the East and the West (Raute 2019). Included studies from Central Europe focus on parental leave and childcare reforms, in general taking place at later time points than reforms from the Nordic context.

The Southern European context

Like the Central European regime, the Southern European regime involves a strong commitment to the traditional family (Ferrera 1996). The Southern European regime includes Greece, Italy, Portugal, and Spain (Gauthier 2002). While health care is tax-financed and provided universally by the states, welfare services are provided in a highly collusive mix among public and nonpublic actors and institutions, and cash subsidies are highly variable (Ferrera 1996). Day care remains scarce, and there is a high proportion of stay-at-home mothers. Scarce social support systems and traditional family

arrangements mean that opportunity costs are high (Hilgeman and Butts 2009). Evidence from this context is likewise scarce but covers transfers, parental leave, and pension reforms.

The Anglo-Saxon context

The liberal regime is characterized by a low level of support for families, targeted at those with greater needs. Benefits are typically modest, giving a strong incentive to work, even for low pay (Esping-Andersen 1990). This regime characterizes the United Kingdom, the United States, and Canada (Gauthier 2002). Fertility policies in the United States, such as the "family cap," have focused on limiting childbearing out of marriage, unintended births or births to young women. The United States lacks adequate publicly funded childcare and has long lacked parental leave schemes (Thomson et al. 2014). When mothers entered the labor market in the 1980s and 1990s, middle class families started using private daycare centers, while poorer families continued to rely on relatives for care (Laughlin 2013). The US system with insurance-based health care, where states are allowed considerable discretion, has given rise to state-specific expansions and cutbacks used to evaluate the effect of the cost of health services on fertility.

In Canada, the governance of health and welfare services falls to the provinces. Quebec is often considered to be quite similar to the Nordic welfare states (Margolis et al. 2021), while the other provinces can more reasonably be described as having a liberal regime (Gauthier 2002). Several policies aimed at supporting families with young children have been implemented in Quebec since the 1980s, and studies of cash transfers and parental leave compensation are included in our review.

Fertility patterns by context

Since the 1970s fertility rates have declined in all of the countries covered in this review. While the decline started early in the Nordic and Anglo-Saxon countries, their fertility rates reversed or stabilized in the mid-1980s to the 1990s, leading to moderately low fertility and relatively high completed fertility in these countries (Sobotka et al. 2020). Increasing fertility at higher ages accounts for the recuperation in the Nordic countries. In the Nordic countries, period total fertility rate (TFR) has fluctuated between 1.7 and 2.3 from the 1970s to 2010, with quite stable cohort fertility rates (CFRs) around 2.0 (Jalovaara et al. 2019). However, since 2010 the TFRs have declined throughout the Nordic countries, despite little change in family policies (Comolli et al. 2021). In Anglo-Saxon countries, recuperation happened despite relatively low public support to families. This has been

attributed to a persistently high number of teenage births (Sleebos 2003), high levels of unintended pregnancies, and very religious subpopulations (Thomson et al. 2014, 491). From 2007 to 2019, the US fertility rate again declined from 2.12 to 1.70 (World Bank 2020).

Fertility rates in Southern and Central Europe declined later, but much faster and to lower levels (Sobotka et al. 2020). Only some countries within Central Europe have seen increasing birth rates more recently. Since the mid-1970s, the average TFR in Central Europe has been stable around 1.5–1.6, with a diverse pattern between single countries ranging from 1.25 in Germany to 1.86 in Belgium. Despite emphasizing family ties for the provision of resources and support, fertility levels in Southern European countries remain substantially lower than in Western and Northern Europe. Since the 1970s to 1980s, the TFRs in Spain and Italy have been "very low" or "lowest low," ranging from around 1.13 to 1.40 (World Bank 2020).

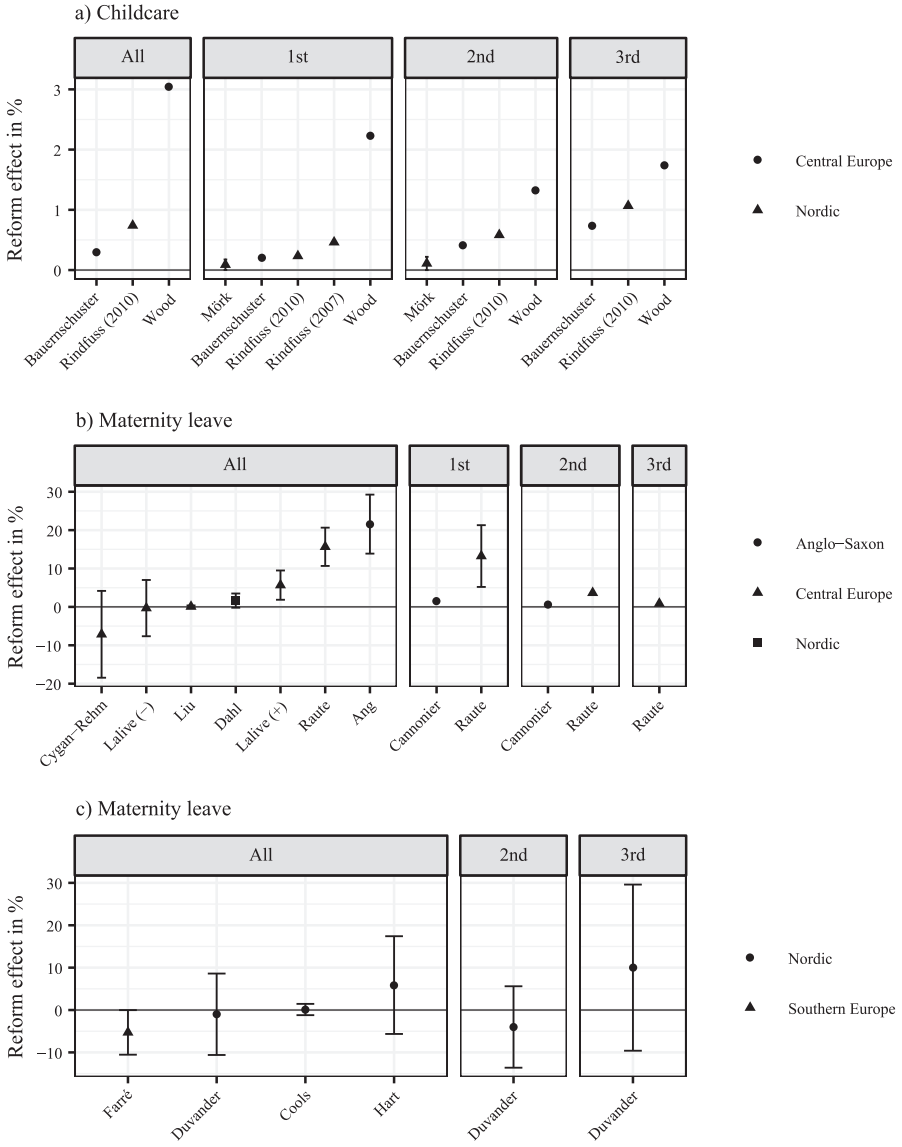
Effect size by policy

Caring for children in and outside the family

Our first group of policies all support families by changing the cost, quality, and nature of childcare. Beyond (often) affecting the family budget, these policies have the potential to affect children's care arrangements. Parental leave polices provide job security so new parents can take time off from work to care for their newborn. If parental leave comes with a prebirth earnings compensation, as opposed to a flat-rate transfer, the value (on an absolute scale) will increase with (maternal) earnings. To the extent that parental leave is taken mainly by mothers, particularly long leaves may cement traditional gender roles. Paternity quotas may compensate for this but could also reduce fertility as fathers' opportunity costs increase. Childcare reforms increase the supply and/or change the cost of formal care, and reduce the conflict of work and family responsibilities for parents (Presser and Baldwin 1980). Childcare reforms have a different social profile, as they also alleviate the care burden for stay-at-home mothers, potentially also increasing fertility in this group.

Public childcare slots increase fertility. Evidence on childcare is drawn from expansions in the Nordic context dating back to the 1970s, and more recent expansions in Central Europe. Effect sizes on relative scale by welfare regime are illustrated in Figure 2a, while details on data and samples are found in Table 2. Note that effect sizes in the figures are recalculated to relative scale and regard the full study sample, while effect sizes in the tables follow what authors report as most important. As such, these need not be identical.

FIGURE 2 Reforms supporting care: (a) Childcare, (b) Maternity leave, (c) Paternity leave



NOTE: Effects of childcare and parental leave reforms on fertility. Results are recalculated to percent (see Data availability statement). When available, 95% confidence intervals are shown with error bars. Estimates preferably for the longest time horizon, and the largest (main) sample or highest age group. Childcare effects are per percentage point increase in coverage. For Mörk, effects are per 1000 Swedish Kronor (~100 USD) reduction in childcare price.

We found solid evidence that expansions of public childcare had positive effects on fertility, both in the Nordic and Central European contexts. The effect of a Norwegian expansion was lasting and substantial; each percentage point increase in childcare coverage for preschool-age children

TABLE 2 Studies of childcare

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Battistin, De Nadai, Padula (2015)[a]	Italian Multiscopo Survey 2009, with fertility reported retrospectively. Main sample is 6,014 married or cohabiting couples, where the woman is aged 23–40 years between 1992 and 2004. Couples with no parent (grandparent) alive are excluded. In total, 5,345 male and 5,840 female grandparents.	Three pension-benefit reforms rolled out 1992–1997 reduced the number of retired grandparents, and therefore potentially the supply of help for (prospective) parents. The reform effect on fertility is analyzed in an RD design.	Age; family tie strength	Number of children of adult offspring. Measured at all parities jointly, one time point, on absolute scale.	One additional grandparent available, increases offspring's fertility by 0.04 children, stronger when family ties are strong.
Bauernschuster, Hener, Rainer (2016)	Administrative data for 1998, 2002, 2006–2009 on childcare and population structure. Register data on all births (women aged 15–44), with information on marital status and parity (for the married).	Expansion of public childcare slots for children under age 3 in West Germany (2005–2008). DiD using childcare coverage in municipality as reform variable. Fixed effects for municipality and year.	Age; marital status; parity (among married)	Births per 1,000 women aged 15–44; age-specific birth rates. Multiple measurements per woman, results on absolute scale.	2.7% increase in birth rates per 10%-point increase in childcare slots, driven by married. Strongest for second and third births and aged 29–33.

/...

TABLE 2 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Mörk, Sjögren, Svaleryd (2013)	Data from Statistics Sweden, Labor Market Policy Evaluation (IFAU), and Swedish Public Employment Service, 1996–2003. Sample is couples with 20- to 45-year-old women, excl. single mothers and unmarried women without children or whose partners are not fathers of their children. Analysis of first births restricted to married couples.	Childcare price reform standardized and capped fees. Announced 1998, implemented 2002. The change in costs varied by household type and across municipalities and is related to change in fertility over time in a DiD design. Time lag to implementation used to identify announcement effects.	Household type (children + income); voting patterns	Child births per 1,000 women aged 20–45. Measured yearly.	9.8% increase in first births in 2000 indicates positive announcement effects, stronger among low-income households. No overall effect on fertility of households with children, but indications of timing.
Rindfuss, Guilkey, Morgan, Kravdal, Guzzo (2007)	Data from Norwegian administrative registers for the years 1973–1998. Birth cohorts 1957–1963 followed from age 15 to 35 (N = 175,722 women). First births only. Aggregate data on childcare supply by municipality.	Explanatory variable is percentage of preschool-age children in childcare centers, assumed to capture variation in supply. Fixed effects for time and municipality (2WFE) net out time trends shared across places, and time-constant municipality characteristics.	Age	Hazard (log-odds) of the probability of a first birth at age *t* (15–35), conditional on no previous births. Yearly measurements, relative scale.	Increase in childcare availability increases first birth risk. Strongest effects below age 25.

/...

TABLE 2 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Rindfuss, Guilkey, Morgan, Kravdal (2010)	Data from Norwegian administrative registers for the years 1973–1998. Birth cohorts 1957–1963 followed from age 15 to 35. Aggregate data on childcare supply by municipality.	Explanatory variable is percentage of preschool-age children in childcare centers. Discrete-time hazard models. Fixed effects for time and municipality (2WFE) net out time trends shared across places, and time-constant municipality characteristics.	Parity (excluding fifth order and above)	Hazard (log-odds) of the probability of each parity transition at age *t* (15–35), conditional on being under risk. Results aggregated into number of children by age 35 using simulation.	10%-point increase in supply of kindergarten slots increases number of children at age 35 with just above a tenth. Positive effect for all parities, strongest for second and third births.
Wood, Neels (2019)	Data from Belgian census 2001 linked to administrative registers (2002–2005). Sample is dual-earner couples (defined in 2001), censored at age 50. N = 157,475 couples at risk of first birth, 216,331 at risk of second, and 321,576 at risk of third.	Explanatory variable is increase in proportion of childcare slots for 0- to 3-year-olds in municipality, utilizing variation over time. Municipality fixed effects net out time-constant municipality characteristics.	Parity	Conditional yearly probability of having a child, estimated in hazard rate models and synthetic parity progression ratios calculated from these.	1%-point increase in local childcare coverage gives 10.8% increase in first birth odds. Smaller effects for second and third births.

[a] Working papers.

increased women's number of children (at age 35) by 0.7 percent (Rindfuss et al. 2010). In relative terms, the strongest effects were found for second-order and higher order births, with a 0.6 and 1.1 percent response per percentage point increase, respectively (Rindfuss et al. 2010).

A comparable German reform found that each percentage point increase in childcare coverage for children below three years of age increased fertility by 0.3 percent overall, concentrated at higher parities (0.2 percent for first births, 0.4 percent for second births, and 0.7 percent for third births) (Bauernschuster et al. 2016). Another analysis from Central Europe compares responses to a Belgian expansion among dual-earner couples and yields substantially larger effects. First birth rates in particular increased by 2.3 percent per percentage point increase in coverage (Wood and Neels 2019). The 1.2 (1.7) percent increase in second (third) birth rates is smaller, but substantially above the German and Norwegian responses. The stronger response at first births suggest effects may (in part) be temporary, but a relatively shorter observation period makes assessment of lasting effects difficult.

Although scaling to effects per percentage point increase in coverage facilitates comparison, it hides the fact that these expansions were large, so that their total effect is substantial. The Norwegian expansion spanned decades and increased childcare supply from 0 to 60 percent, giving a total reform effect of 44 percent. The substantial effects found in these studies may be linked to a high take-up and an analytical design that captures both future and current child effects, so that the estimated reform effect is close to the total reform effect on fertility.

There is also evidence that lower cost of public childcare increases fertility. A Swedish reform, announced in 1998 and implemented in 2002, reduced the cost of public childcare and standardized fees across municipalities. Variation in change by household type and across municipalities was analyzed in a DiD design (Mörk et al. 2013). Among married couples, first births rose by 9.8 percent already in 2000, when the reform bill was passed by parliament but not yet implemented, indicating announcement effects on childless couples. Effects were strongest in low-income households.

Last, we found evidence that grandparental care is important in Southern Europe. Several Italian pension reforms between 1992 and 2001 delayed retirement ages, thereby reducing access to grandparental care, and had negative effects on the fertility of the offspring (Battistin et al. 2015). Formal childcare availability somewhat attenuated these effects, suggesting that effects may be weaker in less familiaristic welfare regimes.

Compensating family care: Longer parental leaves increase fertility. Introductions, changes, and revocations of parental leave are analyzed in seven studies (see Table 3) from Anglo-Saxon, Central European, and Nordic regimes. As sharable parental leave tends to be taken predominantly by the mother,

TABLE 3 Studies of parental leave

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Ang (2015)	Canadian Labor Force Surveys 2000–2008 (monthly representative household survey, rotating panel). Women aged 25–45. Observations collapsed into pre- and post-period. Pre-trends not available.	Quebec Parental Insurance Program (QPIP) in 2006 introduced designated paternity leave, increased max. insurable earnings and income replacement (from 55% to 70% for 30 out of 55 weeks, or: 43 weeks total leave with 75%). DiD with rest of Canada as control.	Parity	Births are proxied by having a child less than one-year old in household, measured yearly. Results on absolute scale.	Positive effect (+23.5%) on overall birth rate. Increased age at birth speaks against tempo effect.
Cannonier (2014)	US panel data 1989–2010 from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). Women with <2 births before 1989.	Family and Medical Leave Act (FMLA) in 1993 introduced 12 weeks unpaid job-protected leave to employees who worked at least 12 months, >1,250 hours at covered employers (depending on company size and sector). DiD with ineligible (working) women as control.	–	Hazard of a first and second birth within the survey period and time to first and second birth.	Increased hazard of first (+5.2% points) and second (+3.0% points) births. Twelve (8.5) months earlier first (second) births.

/...

TABLE 3 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Cools, Fiva, Kirkebøen (2015)	Norwegian register data 1989–2007. Parents of children born three months before and after April 1 in 1991, 1992, and 1993, where both parents had prebirth income above 2× basic amount of the Norwegian social security system.	Introduction of four-week father's quota in 1993. DiD comparing families with children born up to three months before and after reform, relative to the same pre-/post-reform difference of 1991 and 1992. 1992 had a similar reform (+3 weeks of leave), but without father's quota. 1991 as baseline to net out secular trends.	–	Parent's number of children 14 years after reform, and distance to the next child. Results on absolute scale, one post-reform measurement per woman.	No significant effect, point estimate close to zero.
Cygan-Rehm (2016)	The German Microcensus 2000–2012. Women who gave birth Oct.-March 2001/2002 through 2006/2007, 15–45 years old at birth, residing in West Germany.	In 2007, maternity-leave benefits changed from flat means-tested transfers to prebirth earnings compensation (≥67%). Discontinuity design combined with DiD approach comparing mothers with births just before and after the reform. The design exploits variation in benefits for children already born, while benefits for the next child are identical on both sides of the discontinuity.	Employment; eligibility for previous benefit; earnings	Probability of a next birth between 12th and 57th month after a previous birth.	–3.5% points within 33 months, driven by reform losers. Significant negative effects at 45 months. No significant effect after 57 months in the main sample due to reduced precision.

/...

TABLE 3 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Dahl, Løken, Mogstad, Salvanes (2016)	Norwegian register data 1986–2006. Sample is eligible mothers in each reform year (1987–1992) with labor earnings above the substantial gainful activity level in the year before birth.	Six parental leave extensions (total increase 17 weeks, from 18 to 35 from 1977 to 1992). RD for each reform cutoff date for the years 1987–1992. Parents of children born after the cutoff are compared to those giving birth before the cutoff.	–	Number of (additional) children born to a mother 14 years after reform. Results on absolute scale, one post-reform measurement per woman.	1.6% more children after 1992 reform (three-week extension). No significant effects of prior extensions of comparable size.
Duvander, Lappegård, Johansson (2020)	Norwegian and Swedish register data. Parents with children born 25 days before and after each reform, where the birth is the first or second of both parents. Couples divorced or separated are censored. In Norway, only mothers eligible to leave are included.	Introduction of four-week father's quota in 1993 in Norway and 1995 in Sweden. In Norway as extension of the total leave, in Sweden within existing leave length and coinciding with a reduction of earnings replacement (from 90% to 80% for the whole leave, except the father's and mother's quota). DiD comparing pre- and post-reform births in the reform year, using the same periods the year after to net out secular trends.	Parity; income	Yearly odds of having another child within 10 years after reform. Multiple post-reform measurements per woman per outcome.	No general effect. Small positive effect on third births in Sweden if father has low income.

/...

TABLE 3 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Farré, González (2019)	Publicly available microdata and birth-certificate data (2005–2013) from the Spanish National Statistical Institute. Sample is parents having a child around the reform implementation date.	Introduction of two weeks paid paternity leave in 2007 in Spain. RD comparing pre- and post-reform parents. DiD with 2006/2008 parents to address potential seasonality.	Age	Probability of a next birth within two, four, and six years and birth spacing. One post-reform measurement per woman per outcome.	5% reduction in birth probability, driven by mothers >30. Longer spacing between births.
Hart, Andersen, Drange (2019b)[a]	Full population data from Norwegian administrative registers. 2008–2014. Sample is co-resident couples who had a first child between May and August 2009 and where the mother had earnings the previous year.	Extension of the Norwegian father's quota from 6 to 10 weeks on July 1, 2009. The quota compensates 100% of earnings up to a cap and is lost if not taken by father. Subsequent fertility of couples who had a child just before and after this date are compared in an RD design.	Child sex; parity; union type; age; education	Cumulative number of additional children born, measured yearly for five years. Results on absolute scale, multiple post-reform measurements per woman.	No significant effects in main sample, or in any subsample.

/...

TABLE 3 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Lalive, Zweimüller (2009)	Austrian Social Security Database, covering private-sector employees, 1985–2000. Sample is women 15–45 years, eligible for parental leave with births in 1985, 1987, 1990, 1993, and 1996.	Two parental leave reforms: 1990 extension 12–24 months; 1996 revocation 24–18 months (all with 340 EUR flat monthly benefit). RD comparing mothers giving birth a month before and after each reform (current-child effects). Revocation of reform after three years and 1987 mothers used to identify future child effects. Robustness checks include pre-/post-trends (1985/1993 mothers) and ineligible women.	Income (above/below median); occupation (white/blue collar)	Probability of another child within three years, and within 10 years for 1990 reform. Future child effects measured within three-year window (birth probability within six years as control). Several shorter intervals used to detect timing.	Extension for the current child increased birth prob. within 10 years by 5.7%. Strongest effects within three years (+14%) and among women with low income. Extension for future children increased birth prob. within three years by 24%. Reduction affected timing only. Timing in line with incentives.
Liu, Skans (2010)	Data from population-wide registers (1986–2005) including all children born in Sweden between March and December 1985, 1987, 1988, and 1989.	Parental leave extension (12–15 months) in 1989, with retroactive coverage for births Oct. 1988 onward, and extension of 30 days (60 days) for births in Aug. (Sept.) 1988. DiD using variation in entitlement between March to July 1988 (pre-) and Aug. to Dec 1988 (post-). 1987/1989 births used to net out secular trends.	Education	Probability of having an additional child within 18, 24, 36, and 72 months, and total number of children.	No effect in the main sample after 72 months. Short-term effect (within 18 months) of +4.1% per extra month of leave entitlement. Effects driven by highly educated mothers.

/...

TABLE 3 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Raute (2019)	German Pension Registry, 2004–2012, using 1% of population of insured (economically active) women aged 21–44, for comparison by earnings. German Microcensus, 2003–2012, 1% of nationally representative population aged 21–44 (econ. active and inactive) for comparison by education.	In 2007, maternity leave benefits changed from flat means-tested transfers (300–450 EUR monthly for low-income households max. 12–24 months), to prebirth earnings compensation of at least 67% for max. 12 months. DiD using women below 35th income-percentile (alt. lower educated) as control. Importantly, the reform effect identified in this design causes persistent variation in benefits received for future children.	Age; parity	Yearly birth probability in a given year (first and higher order births), scaled to births per 1,000 women. Multiple post-reform measurements per woman.	+16% increase in main specification (comparing across income percentiles). Stronger effect if treatment group is based on educational level. Effects on first and second births at high ages indicate effects on completed fertility.

[a] Working papers.

this discussion occurs in conjunction with maternity leave. Reform effects as percentage change with 95 percent confidence intervals from each study's main sample are plotted in Figure 2b.

Large parental leave reforms in Central Europe had significant and substantial effects on fertility. Here, public childcare is scarce, particularly for younger children, so that paid parental leave will replace unpaid leave. An Austrian reform in 1990 doubled the leave period from 12 to 24 months, yielding a 5.7 (14) percent higher likelihood of another birth within 10 (3) years among mothers who were entitled to longer leave (Lalive and Zweimüller 2009). A revocation from 24 to 18 months in 1996 had no effect on births within three years (the longest time horizon observed), though births were shifted to an earlier time point within this time frame.

In 2007, German maternity leave benefits changed from being flat, means-tested transfers available for up to 24 months after birth, to compensate prebirth earnings for up to 12 months. Comparing the probability of another birth for mothers with a birth shortly before and after the reform, Cygan-Rehm (2016) found a statistically insignificant decrease in births in the main sample after 57 months. Among low-earning mothers, who were worse off as a result of the loss of the longer flat rate benefit, the reform significantly reduced the probability of having another child within 21–45 months. Raute (2019) analyzed the same reform, contrasting outcomes for long-term "winners" (higher educated/earning) with long term "losers" (lower educated/earning). For these groups, benefit levels vary, even for future children and first births, so the estimated reform effect captures more of the total reform effect on the target group. This likely explains why the estimated reform effects are substantially larger—a 16 percent increase in the yearly birth probability concentrated at the first and second parity. Effects are measured for up to five years after the reform, but large effects even in the highest age groups indicate that the impact is not transitory.

From the Anglo-Saxon context, the United States' introduction of 12 weeks of unpaid job-protected parental leave in 1993 increased the probability of first and second births by 1.5 and 0.6 percent annually, respectively (Cannonier 2014). Effects are identified by comparing change over time in birth rates among women eligible and ineligible for job protection. This is a large effect of a relatively small expansion. Possibly, effects are larger for expansions from short or nonexistent leaves, as there are few good substitutes for parental care for newborns. A 2006 reform in Quebec introduced a seven-week paternity quota and increased parental leave benefits from 55 to 70 percent earnings replacement for 30 (of 55) weeks. Ang (2015) estimates that this reform increased the birth rate in Quebec by 23.5 percent compared to other Canadian provinces. In absolute terms, effects were

strongest for second and fifth births, but differences by parity are not statistically significant and may simply be chance. The absence of robustness and pre-trend tests, and the many other Quebec-specific family policy reforms, calls for some caution in interpreting these effects.

In the Nordic context, Dahl et al. (2016) analyze a series of six small (three to four weeks) expansions in parental leave in Norway, comparing women who gave birth just before and just after each expansion. Only one expansion had a marginally significant impact on mothers' number of children 14 years later (+1.6 percent, $p < 0.1$). An expansion of the Swedish paid parental leave from 12 to 15 months similarly had a quite precisely estimated zero effect in the main sample (0.2 percent increase in total number of subsequent children, $p > 0.1$) and temporary positive effects on highly educated mothers' fertility (Liu and Skans 2010). These two policy reforms compare outcomes for women who differ in their experience with leave for current children but have the same prospects for leave for any future child. Along with no variation in benefits for future children, the two Nordic studies analyze small reform increments, making it likely that the estimated reform effect will be substantially smaller than the total reform effect.

Differences in reforms and analytical design by study context mean that evidence is insufficient to conclude whether effects are context-dependent. However, when studies using large changes and capturing future child effects are given most weight, it seems clear that longer parental leaves increase fertility, both in the short and long term.

Paternity quotas: Ambiguous predictions from theory and insufficient evidence. Paternity quotas were introduced as a response to mothers taking most parental leave, aiming to ensure less traditional gender patterns in care and paid work. Paternity quotas are established by reserving sharable weeks for the father, and/or by adding weeks reserved for the father to the total parental leave. Evaluations of the introduction and extensions of paternity quotas in the Nordic countries show no significant effects in the main samples (see Cools et al. 2015 for its introduction in Norway; Duvander et al. 2020 for its introduction in Sweden and Norway; and Hart et al. 2019b for the extension in Norway). This holds both in the short and in the long run (see Table 3). Each incremental change in the paternity quota is small, and the analytical designs applied only identify current child effects. Although their experiences differ, treatment and control groups of parents will have the same paternity quota if they have another child. As such, these analytical designs capture only a small portion of the total reform effects, and the absence of effects should not be taken as evidence for no effect.

Evidence from the Southern European context comes from Spain, where a 13-day fully compensated paternity leave period was introduced

in 2007 (Farré and González 2019). This was in response to the observation that the 10 weeks of parental leave available to share from 1999 was taken near exclusively by mothers. Using an RD design, Farré and González (2019) found that the paternity quota reduced fertility by 5 percent. The direction of the effect is in line with increased opportunity costs for fathers, and the magnitude of the effect is substantial given a relatively small change in opportunity costs.

Reducing the monetary cost of childbearing

A second broad group of policies reduce mainly the monetary cost of childrearing. Without public health care, perinatal care and health services for the new child constitute a large proportion of the immediate costs of having a child. Public health services and cash transfers to families with children reduce the costs of current and future children and may have positive effects on fertility. Increasing financial transfers may also shift the cost of raising each child permanently upward (cf. the notion of "child quality," Becker 1991). As such, transfers may impact the manner in which families live, rather than their size. The monetary value of cash transfers tends to be smaller than that of public childcare and compensated parental leave (Thévenon and Gauthier 2011).

Universal transfers increase fertility. Evidence on transfers is taken from all four welfare regimes. All but one Spanish reform took place in the previous century (Table 4). Effect sizes are plotted in Figure 3a.

Studies from Quebec, with supporting evidence from the Nordic context, suggest a positive yet potentially transitory effect of transfers on fertility. A substantial increase in transfers, particularly at third births, was implemented in Quebec in 1988. Milligan (2005) estimates a 12 percent increase in fertility following the expansion. The effect is the strongest for third births, where the increase was the largest. Effects at higher order births and higher ages indicate an impact on quantum, but with an observation period of five years, conclusions remain tentative. Using a comparable design and time frame, Ang (2015) suggests only a 1.72 percent increase. We put more emphasis on the more credible design of Milligan (2005).[3] Parent and Wang (2007) explore permanent effects of transfers further, by estimating short- (5-year) and long-run (15-year) effects of a change in the same subsidy in the 1970s. While the short-term increase in fertility is of the same magnitude as found by Milligan (2005), long-term estimates suggest that a relative fall in birth rates follows, leaving cohort fertility unaffected (Parent and Wang 2007). A smaller regional increase in transfers to families with children was linked to a 4 percent increase in fertility in Northern Norway (Galloway and Hart 2015). The effect is found only among women in their

TABLE 4 Studies on universal child transfers

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcomes	Reform effects
Ang (2015)	20% sample from Canadian Census. Households with a woman aged 25–45 who is a Canadian citizen. Measures taken in 1986 (pre-), 1991, and 1996 (both post-). Pre-trends measured at aggregated level only.	Reform is increases in cash transfers in Quebec. Most importantly, Allowance for Newborn Children (ANC) increased substantially in the period 1987–1992. DiD design with Quebec as treated group and rest of Canada as control.	Parity	Number of children 5 and 10 years after the reform, based on age and parity of children from Census records.	+1.72% significant effect in main sample.[a] Effects concentrated at first and 5th+ parities.
Galloway, Hart (2015)[b]	Data from Norwegian administrative registers. Sample is all citizen women residing in the county Troms at least one year 1984–1988. Person-years for age <15 and >39 are dropped. Fertility outcomes observed until 1997.	Increased regional cash transfer (+$350 in 2019 USD) in Northern Troms, Norway (1989–1990). Analyzed in DiD design with Southern Troms as control.	Age, parity, marital status	Yearly probability of having a(nother) child, averaged over a nine-year reform period.	+4% ($p < 0.05$) among women aged 20–24 (joint analysis by age not shown). Effects concentrated among unmarried.
Gathmann, Sass (2018)	German Microcensus (annual cross-sectional survey on 1% sample of the population) 2000–2010. Study sample is East German households with at least one two-year-old child.	In 2006, the East German state Thuringia introduced a monthly "home-care subsidy" for families who did not send their two-year-old to public childcare (€150 for firstborn, up to €300 if fourth child). DiD comparing Thuringian families with eligible children to those in the rest of East Germany before and after policy change.	Parity, family status, education, income, citizenship	Probability of having another child (proxied by having a child under 12 months in the household).	No general effect on subsequent births; small pos. effects if 2+ children; stronger among single mothers, low-income, and foreign parents.

/...

TABLE 4 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcomes	Reform effects
González (2013)	Vital statistics with time of birth (by month) and length of gestation for the full population of Spain from the Spanish Statistical Institute for the period 2000–2009.	2,500 EUR baby bonus, paid immediately at birth to mothers with at least two years residence, announced in Spain July 2007, taking effect immediately. Change in birth rates around announcement analyzed in an RD design.	None	Monthly number of conceptions (multiple measurements in post period), backdated using birth month and gestation length.	+5% ($p < 0.05$) in main sample.
Groves, Lopoo (2018)	US National Longitudinal Survey of Youth 1979 (aged 14–22 in 1979). Attrition before 1988 excluded. Treatment group is youth who had a father who died before the age of 18, control group not. Cohorts set to graduate from high school 1982 or earlier were eligible for the benefit (if father dead), younger cohorts were not and constituted a control group. Outcomes are measured in 2012, or the latest wave the resp. is observed.	Discontinuation of subsidies (~$6,700 in 1980 USD per year) to unmarried students <23 years with a disabled, retired or deceased parent in the United States (1982–1985). DiD design, first difference between lost parent versus not (graduated 1982 or before), the second between lost parent versus not (graduated 1983 or later).	Gender	Ever had children, and age at birth (if ever children). One measurement per person, absolute scale.	No significant effect on propensity to ever have children ($p > 0.1$). Significant increase in age at birth.

/...

TABLE 4 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcomes	Reform effects
Milligan (2005)	Canada/Quebec: Vital statistics and Canadian 1991 and 1996 Census Public Use Microdata Files on Families (PUMF-F). Alternative data sources produced results supportive of the Census results, but with less precision.	Increase in ANC in Quebec (see Ang 2015 above). Effects estimated in a DiD design comparing Quebec to the rest of Canada, and also in a DiDiD design using women at lower parities (where the increase was less substantial) as a second control group.	Parity, family income, mother's education	Probability of having a child between survey waves, measured jointly for all parities and separately by parity. One measurement per woman in post period, absolute scale.	Overall reform effect is +12% ($p < 0.05$) prob. in having a child in main sample. Effects concentrated at higher parities and in higher income brackets.
Parent, Wang (2007)	Canada/Quebec: Canadian Public Use Microdata Files on Families (PUMF-F) and Individuals (PUMF-I) 1976 and 1981 ("short term"), 1986, and 1991 ("long term"), and vital statistics. Resident families where female is aged 25–39 and had not moved province five years before the census time.	Quebec-specific expansions of Family Allowances Program in Canada in 1974. Benefits increased from C$60 to C$240 and varied with number of children and spacing. DiDiD comparing the development in Quebec with the rest of Canada, using lower parities as a second control group.	Age, parity	Whether there are children aged <6 present at home. Measured in the short (5 years) and long (10 years) term.	No significant long-term effects. +10.1% in short term, strongest for higher parities.

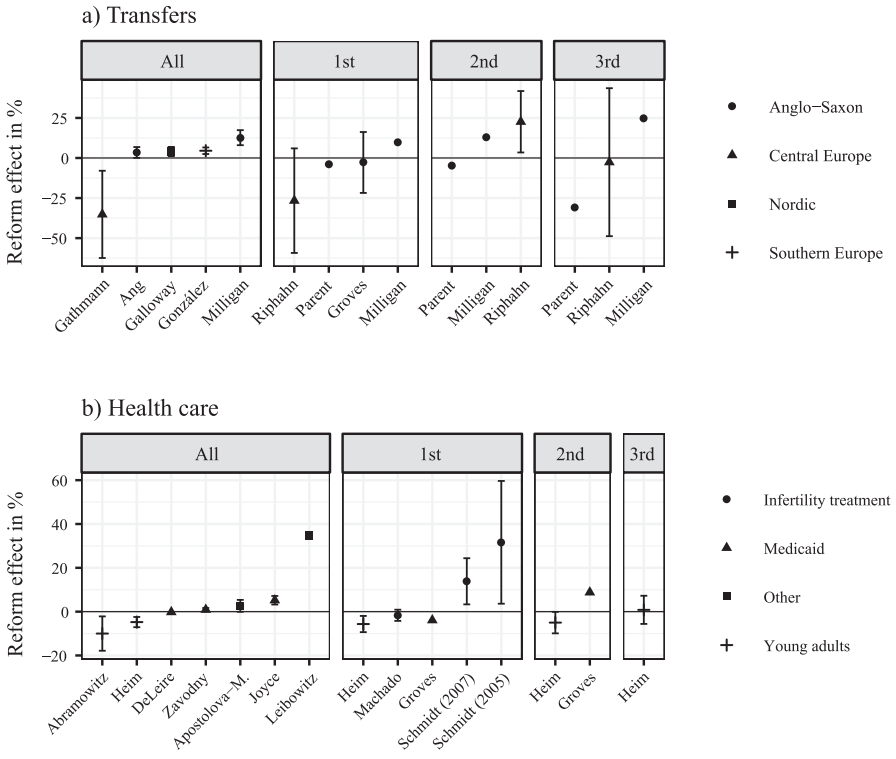
/...

TABLE 4 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcomes	Reform effects
Riphahn, Wiynck (2017)	Longitudinal data from SOEP (German Socio-Economic panel) 1992–1998 waves and Microcensus 1995–1999 waves. Samples include married and nonmarried cohabiting couples with woman aged 25–40, no welfare recipients, both German residents living in West Germany. Two years pre-reform and two years post-reform is measured.	Pre-reform, Germany combined income-dependent cash transfers plus tax-exempt child allowances at birth. The 1996 reform significantly increased the cash benefit. Households choose between the tax deduction and cash benefit; which pays best varies with income and by parity. These heterogenous reform effects are used in a DiD design, where couples with low versus high educational attainment alternate as treatment and control group for different parities.	Education, parity	Propensity to have a first child and second child, measured yearly on an absolute scale. Constructed from info on whether the couple has a child under the age of 1 living in the household.	Negative effect on first births among lower educated ($p < 0.1$), contrary to incentive change. Positive effect on second births among higher educated ($p < 0.05$).

[a] Descriptive statistics for births in pre-period in treatment group are not given. The percentage change is the author's own scaling relative to the overall trend in Quebec (Ang 2015, 280).
[b] Working papers.

FIGURE 3 Reforms reducing monetary costs: (a) Transfers, (b) Health care



NOTE: Effects of transfers and health service reforms on fertility. Results are recalculated to percent (see Data availability statement). When available, 95% confidence intervals are shown with error bars. Estimates preferably for the longest time horizon, and the largest (main) sample or highest age group.

early 20s and concentrated at first birth, suggesting a tempo shift rather than an effect on completed fertility.

From the Central European context, Riphahn and Wiynck (2017) exploit a German reform in 1996. The reform increased subsidies for first births among low-earning couples; at this parity high-earning women saw no change and are the control group. Surprisingly, birth rates fell among low-earning relative to high-earning women ($p < 0.1$). For second births, benefits increased among high-earning couples, while low-earning women had no change and are the control group. Here, birth rates increased relatively more for high-earning couples, in line with the incentives ($p < 0.05$). A 2006 reform in the East German state of Thuringia transferred at least 150 euros monthly (more to larger families) to those not sending their two-year-old child to public childcare. Comparing responses in Thuringia to other East German states, Gathmann and Sass (2018) found positive effects on higher order births, measured in the first four years following

the reform. The effects were concentrated among groups more prone to home care (large families, single mothers, low-income, and foreign parents).

In regimes with less extensive state support to families, transfers have a positive effect on fertility, too. In 2007, a sizeable universal child benefit was introduced with the explicit aim of counteracting falling fertility rates in Spain. This introduction was linked to an immediate 5 percent increase in fertility (González 2013). Although this effect is stable through 2.5 years observed after the reform, the time horizon is too short to distinguish effects on quantum and timing. The reform increased mother's time out of the labor force and reduced children's time in formal care. In the Nordic context, no effects on maternal labor supply were found (Galloway and Hart 2015).

Finally, Groves and Lopoo (2018) analyze the retraction of federal aid to US students with one deceased parent if the students had children themselves. The subsidy worked as a disincentive to early childbearing only, and significantly increased age at first birth. They find insignificant (and imprecisely estimated) effects on the probability of ever having children. Long-term effects would be a surprising response to an incentive to postpone childbearing in the early 20s.

Overall, there is quite solid evidence that increases in transfers have an immediate effect on fertility, observed in all four welfare regimes. Most of the designs are unable to test whether effects are lasting or not, and some provide evidence that effects are transitory. We note that the regional designs used in most studies of transfers capture benefit changes for future children well but have some challenges with measurement error for long-term outcomes.[4]

Health care access facilitates intended births and reduces unintended births. The insurance-based health system in the United States, with substantial variation over time at the regional (state) level, facilitates identification of effects of access to health care on fertility. All 10 studies are based on US reforms and data. Cheaper health services reduce both the cost of a child and regulation costs, and their effects depend strongly on age group. All studies are described in Table 5, and effect sizes are plotted in Figure 3b.

The largest positive effects on fertility are found for the provision of infertility treatment. For women nearing the end of their fecund period, cheaper infertility treatment increased fertility (Schmidt 2005, 2007). Insurance-covered infertility treatment increased birth rates by 32 percent among (white) women above the age of 35 (Schmidt 2005). In contrast, Machado and Sanz-de-Galdeano (2015) found no effects on the probability of having at least one child by age 35, when the need for fertility treatment is substantially lower.

TABLE 5 Studies on health services

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Abramowitz (2018)	American Community Survey data (ACS) 2008–2010 and 2012–2013. Women 20–25 (629,300) and 28–30 (319,600) from the United States, living in households or college dormitories. Active duty military are excluded.	The Affordable Care Act (ACA), implemented in the United States 2010, allowed young adults to be covered on their parent's insurance until they turn 26, reducing their cost of contraception, birth, and abortion. Change in fertility among eligible young adults (aged 19–25) was compared to change in fertility among older (28–30) young adults in a DiD design. Pre-trend plots and tests.	Age, marital status, race, urban/rural	Yearly probability of having given birth within last 12 months. Multiple measurements per woman in reform period, absolute scale.	10% ($p < 0.01$) reduction in probability of having given birth,[a] concentrated among whites.
Apostolova-Mihaylova, Yelowitz (2018)	American Community Survey data (ACS) 2003–2011, excluding 2007. The full sample is 507,000 women 15–44 in Massachusetts and surrounding states. Subsample: 134,910 women are included if they resided in New England and do not have imputed values for gender, fertility, age, marital status, or race.	Massachusetts health care law of 2006 mandated employers to provide health insurance, lowering the cost of pregnancy and contraception. Change in fertility studied in DiD design, with women from neighboring states as control.	Age, marital status	Yearly probability of birth in the last 12 months. Multiple measurements per woman in reform period, absolute scale.	No effect in main sample. 1% increase among married women aged 20–34; 8% decrease among unmarried.

/...

TABLE 5 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Deleire, Lopoo, Simon (2011)	Full population US birth data from the National Center for Health Statistics, 1985–1997, with information on race, state of residence, education, and age. Study sample is 53,856 white and African American women aged 15–44. Collapsed to 44 cells by state quarter. Fraction eligible by cell and state year calculated using information on income and marital status from March Current Population Survey (1991–1994).	Expansion of Medicaid availability in the United States 1985–1996 (measured as the simulated fraction available), using variation over time and between states in a 2WFE design, also controlling for demographic characteristics.	Race, marital status, education, age	Quarterly birth rate (logged), by cohort/race/state/year.	No significant effect in main sample after controls (–0.2% per 1 percentage point increase in fraction available for Medicaid, $p > 0.01$).
Groves, Hamersma, Lopoo (2018)	Birth counts from Vital Statistics (1987–1997) for all whites and African Americans. Additionally, age, parity, education, and marital status define cells. Cell-specific population counts from 1980 to 2000 Public Use Microsamples (PUMS). Collapsed by cell, state, and quarter/year.	2WFE design that utilizes change in the eligibility for public health insurance across states and time. Reform variable is threshold to have the right to public health insurance coverage, as a fraction of the federal poverty limit. Extensive checks, incl. limitation to federally initiated changes to avoid self-selection.	Race, marital status, education, age, parity	Quarterly birth rate (logged), by cohort/race/state/year.	9 (7)% ($p < 0.001$) increase in higher order births for African Americans (whites). Concentrated among high-school educated women.

/...

TABLE 5 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Heim, Lurie, Simon (2018)	US federal tax documents and Social Security Administration (SSA) 2008–2013. 401,922 young adult women 24–29, excluding 26-year-olds. 1% sample from the parent–child matched population.	Introduction of ACA in 2010 (see Abramowitz above). DiD design where young adults aged 24–25 are treated, and 27- to 29-year-olds are the control group.	Parents' insurance, parental income, sex, marital status, parity, post-secondary enrollment	Conception resulting in live birth. Measured yearly, results on absolute scale.	5% ($p < 0.01$) reduction in fertility in main sample, concentrated among unmarried and lower educated women.
Joyce, Kaestner, Kwan (1998)	National Center for Health Statistics (NCHS) data on pregnancies for 15 US states 1987–1991. Data on births 1986–1992 from National Natality files. Sample is unmarried, lower educated women aged 19–27 with 12 or fewer years of schooling. Population size from PUMS 1990 (5% sample).	State-specific expansions of Medicaid insurance for pregnancy, perinatal, and child health analyzed in a 2WFE design with state and year fixed effects. Threshold was raised in two steps. Each stepwise raise is indicated by a dummy (these are the reform variables).	Race	Quarterly birth rate (logged), by cohort/race/state/year.	5% ($p < 0.001$) increase in births rates among whites as response to the first expansion, comparable to the second expansion. No effects among African Americans.

/...

TABLE 5 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Leibowitz (1990)	Experiment participants were chosen randomly from six US cities in Washington, Massachusetts, and South Carolina. Sample is 2,216 participants aged 15–44. Participants in Seattle were censored if they moved out of state.	The RAND Health Insurance Experiment (HIE) randomly assigned families to different insurance schemes, creating random variation in the cost of health care. The study tests the effect of free medical care on fertility over three to five years (1974–1979), using those assigned to shared costs as a control group.	None	Number of pregnancies leading to live birth/stillbirth over the experiment period. Recalculated to births per 1,000 and annual pregnancy rates per 1,000.	29% ($p < 0.05$) increase in births in main sample. Annual pregnancy rates show strongest increase after two to three years.
Machado, Sanz-De-Galdeano (2015)	Microdata on births from June Current Population Survey (CPS) used for individual-level analysis.	Expansions in mandates for insurance to cover infertility treatment. Variation across states and over time studied in a 2WFE/DiD design. Effects vary with time since enactment.	Race	Probability of having at least one child by age 35 (30). One measurement in reform period, absolute scale.	No overall effect on prob. of having a child by age 35. Some postponement effects, that is, negative effects at age 30.

/...

TABLE 5 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Schmidt (2005)	First birth counts from Vital Statistics Detail Natality Data and population estimates from Census Bureau. Observation period is 1985–1999; 15 treatment states. Data set aggregated by cohort/race/state/year-cell, dividing number of first births by population size.	Variation between states over time in whether infertility treatment is included in insurance. DiDiD-design controlling for state and time fixed effects, with mandate to cover infertility treatment as reform variable, interacted with age >35 (main users of infertility treatment).	Race	First birth rate (logged), by cohort/race/state/year.	32% ($p < 0.05$) increase among women over 35, concentrated among whites. No effects for younger ages.
Schmidt (2007)	First birth counts from Vital Statistics Detail Natality Data and population estimates from Census Bureau. Observation period is 1985–1999; 15 treatment states. Data set aggregated by cohort/race/state/year-cell, dividing number of first births by population size.	Expands upon Schmidt (2005), additionally distinguishes between strong or weak mandate; IVF covered or not; covered proportion of population. Robustness tests include state-specific trends and restricted time series.	Race	First birth rate (logged), by cohort/race/state/year.	22% ($p < 0.05$) increase in first births among whites. Effect increases with (but not dependent on) mandate strength and if a larger proportion is covered.

/...

TABLE 5 (continued)

Authors	Data and sample	Reform and identification strategy	Separate analysis by	Outcome	Reform effects
Zavodny, Bitler (2010)	Panel data of US women aged 15–44 for the period 1982–1996. Birth rates aggregated by cell, state, and quarter/year. Numerator is birth counts from National Center of Health Statistics (NCHS). Denominator is constructed using population counts from the Census Bureau, and distribution by covariates from the March CPS.	Effect of changes in Medicaid availability (simulated fraction eligible and income threshold). Women are eligible for Medicaid if earnings are below threshold. 2WFE netting out state and year fixed effects.	Race, marital status, education	Ln(quarterly birth rates), aggregated by cell/state/year/quarter. Results on relative scale.	1% increase in birth rates when eligibility threshold increases 100% among white women, concentrated among the lower educated. No overall effect of fraction eligible in the main sample.

[a] Scaled to the birth rate in age group 20–25, as suggested by authors.

In the youngest age groups, access to health insurance reduced fertility, likely as a result of reduced regulation costs. The Affordable Care Act ("Obamacare," ACA) of 2010 reduced the cost of health care for young adults, as they were entitled to be included on their parents' insurance. Studies using unaffected age groups as control groups suggest a reduction in yearly birth rates between 5 percent (Heim et al. 2018) and 10 percent (Abramowitz 2018). The age profile and mechanism of regulation costs suggest that the effects affect timing only.

In contrast to these substantial effects, quasi-experimental evidence suggests no or very small effects of state-specific extensions of Medicaid. Medicaid is a means-tested public insurance, covering a large share of the costs of perinatal care, delivery, and health services for low-income US families. A simple 2WFE design suggested a 5 percent increase in fertility, albeit insignificant in the main sample ($p > 0.1$) (Joyce et al. 1998). Designs with more detailed controls suggest a 0.2 percent reduction in fertility ($p < 0.1$) (Deleire et al. 2011), or a precisely estimated but minuscule 0.9 percent increase in fertility from a substantial 100 percent increase in the eligibility threshold ($p < 0.05$) (Zavodny and Bitler 2010). It is possible that the socioeconomic groups affected by Medicaid changes respond less to monetary incentives than the groups affected by ACA, as was found for transfers (Milligan 2005).

Two studies assess the effect of health services on a broad range of ages and socioeconomic groups. First, six US cities randomly assigned families to different health insurance schemes in 1974–1979. Those assigned to a fully covered plan with free health services had a 29 percent higher birth rate than the control group (Leibowitz 1990). These effects may have been a transitory response to the experiment; if people are aware that the cost of health care will increase later, they may pace up childbearing. Second, a predecessor to the ACA was rolled out in Massachusetts in 2006, yielding 8 percent lower fertility among unmarried women aged 20–34, a group for whom births are often unplanned (Apostolova-Mihaylova and Yelowitz 2018). Among married women in the same age group, who often intend to have a(nother) child soon, fertility increased by 1 percent, comparable to the Medicaid effect sizes.

Synthesizing the evidence

Policies that try to ease economic and work–family constraints might play out differently by welfare regime and likewise affect different population groups within the same welfare state differently (Cherlin 2016). In this section, we discuss how policy effects vary by welfare regime and population group, and potential sources of selectivity and bias in the literature. Overarching conclusions by policy type are summarized in Box 2.

Box 2
Which policies affect fertility?

Positive effects of childcare expansions on fertility are documented in the Nordic and Central European contexts. Effects are largest in the short run, but substantial also for (near-)completed fertility. There is no contrarian evidence, but evidence is lacking in the Anglo-Saxon and Southern European contexts where reliance on private care providers and relatives is high.

Substantial expansions of parental leave schemes increase fertility. This conclusion is based on evidence mainly from the Central European and Anglo-Saxon (US and Quebec) contexts. In the Nordic context, expansions happened earlier and in smaller increments, and the reform designs therefore rarely capture "critical junctures" (cf. Neyer and Andersson 2008).

Transfers have a positive effect on fertility, but the effect is likely transitory. Evidence is drawn from all four welfare regimes. Cutbacks in welfare in Anglo-Saxon regimes seem to have little or no effect on fertility, but these are of a quite different nature than transfers (see Appendix A in the Supporting Information).

Cheaper health services in general and cheaper assisted reproduction in particular have a positive effect on fertility. This conclusion is based solely on findings from the United States, and whether effects will be similar in other contexts is an empirical question.

For childcare and parental leave, effects are not only temporary. Effects are more likely to be permanent effects if they are measured over several years, occur toward the end of the reproductive period, or at higher parities. For transfers, there is some evidence that effects are waning over time.

Effects by welfare regime

The Nordic and Central European regimes: Convergence. Our review supports a notion of a "Nordic fertility regime" (cf. Andersson et al. 2009), where policies that help families balance work and care contribute to high fertility. From both the Central European and Nordic regimes, we find solid evidence of positive effects of childcare (Bauernschuster et al. 2016; Mörk et al. 2013; Rindfuss et al. 2007, 2010; Wood and Neels 2019). Features of parental leave reforms in Nordic countries mean that the estimated reform effect is a poor measure of the total policy effect (Cools et al. 2015; Dahl et al. 2016; Duvander et al. 2020; Hart et al. 2019b; Liu and Skans 2010). For parental leave, studies from Central Europe capture more substantial policy changes, and bear evidence of positive fertility effects of substantial magnitude (Lalive and Zweimüller 2009; Raute 2019).

Taking evidence from the Nordic and Central European contexts together, parental leave and childcare both seem to have the potential to shift a society to a higher fertility equilibrium (cf. Esping-Andersen and Billari 2015). The notion that a certain "package" of family policies must be in place for them to take effect has been pointed out previously (Gauthier 2007; McDonald 2006) but is not striking from our evidence. Rather, we see large effects of single policy changes in the Central European context, where support to families in general is scarcer. For large reforms to have an effect, no "package" is needed; on theirs own they can already be substantial enough to affect fertility decisions.

The Southern European regime: A traditional equilibrium difficult to shift. In the Southern European regime, some relatively scattered reforms have been implemented in a context of sustained low fertility, traditional family values, and weak public support to families. Transfers had positive effects in Spain as in other welfare regimes (González 2013). This could be because transfers in general are less context dependent, but also because economic concerns were particularly pressing in countries hit hard by the 2009 recession. The introduction of the paternity quota, aiming to de-traditionalize care patterns, reduced fertility in Spain (Farré and González 2019), perhaps unsurprising in light of high dependency on male breadwinning. An Italian reform aiming to increase the retirement age had an unintended negative effect on fertility (Battistin et al. 2015), because some Italian parents had to rely on grandparents in the absence of formal care. These effects are powerful illustrations of how a decontextualized and incoherent approach to family policies might fail, when single policies are implemented without considering the predominant care arrangements in the respective society. The contrast between reform effects in Central and Southern Europe constitutes a warning for policymakers; if (relatively) high fertility is the goal, *ad hoc* measures cannot replace reliable policies that support parents in caring for their children.

The Anglo-Saxon regime: Limited family support and limited evidence. Polices supporting families, and studies that analyze them, are generally scarce for the Anglo-Saxon regime. One study finds positive effects of job-protected parental leave (Cannonier 2014), comparable to effects found in regimes with more support to families; again, this is counter to the expectation that reforms are more efficient when provided in a "package." Although evidence of policy effects is abundant from the Canadian region Quebec, its extensive family support policies are atypical for the liberal welfare regime. The external validity of the studies from Quebec with regard to the rest of the liberal regime is therefore questionable.

The comparatively large share of unintended births in the United States has received substantial attention. By reducing unplanned fertility in the early 20s and increasing (intended) births toward the end of the

fecund period, cheaper health services seem to restructure when in the life course childbearing happens. The polarized fertility pattern in the United States may to some extent be the result of a lack of affordable health services and thereby high regulation costs in the lower income population. Welfare cutbacks, on the other hand, have no or very small effects on fertility (see Appendix A in the Supporting Information), but the nature of these reforms, and the subpopulation they cater to, means that they say little of the potential effect of more classical family policies in the United States.

Amplifying inequality or leveling the playing field?

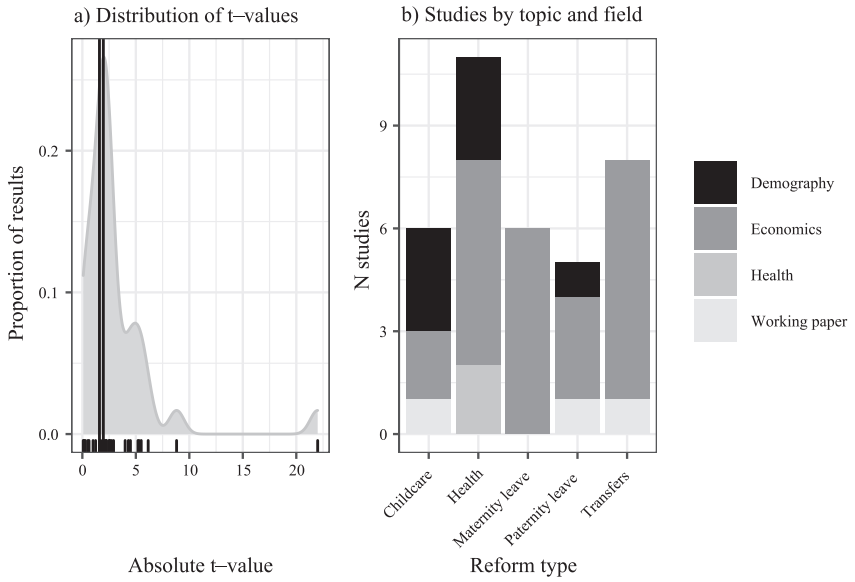
Mothers' employment and care burden differ across social groups (McLanahan 2004). Policies that enable mothers' employment and protect their careers should expectedly be more important for women who have invested in higher education. Economic constraints and financial transfers, on the other hand, should matter more for fertility decisions in households with lower income (potential).

Parental leave often has the goal of securing mothers' labor attachment and in our evidence from parental leave reforms women's education and/or income was considered important in more than half of the studies. Two studies of change in the parental leave benefits in Germany illustrate diverging responses by women's socioeconomic status: the loss of a long flat-transfer leave reduced births among low-earning women in the short run (Cygan-Rehm 2016), while high-earning/higher educated women saw a relative increase in births with earnings-compensated parental leave (Raute 2019). In Austria, too, flat-rate transfers had more effect on low-wage women, while women both in white- and blue-collar occupations responded positively to extended job protection (Lalive and Zweimüller 2009). The extension of earnings-compensated parental leave in Sweden had strongest short-term effects on highly educated mothers (Liu and Skans 2010).

Parents' socioeconomic background was no major stratification variable for childcare (used in one of six studies). We note that the only study on childcare coverage that restricted the sample to dual-earner couples found relatively weaker effects on higher parities but far stronger effects on first births (Wood and Neels 2019), as opposed to coverage studies that included all women of childbearing ages (Bauernschuster et al. 2016; Rindfuss et al. 2010). Price changes introduced in the Swedish childcare reform led to differing costs according to household income, with stronger fertility responses in low-income households (Mörk et al. 2013).

In line with expectations, cheaper health services had stronger effects on the fertility of lower- or high-school educated women (4 of 11 studies were stratified). When results for transfers were stratified by income or education (three of eight studies), effects were strongest in the higher

FIGURE 4 Bias assessment: (a) Distribution of *t*-values, (b) Studies by topic and field



NOTE: In Panel a, density plot shows smoothed distribution of t–values. T–values are calculated from the main model and main sample of each study included. Vertical lines indicate the thresholds for p<0.1 (t=1.6) and p<0.05 (t=1.96). Rug plot along the x–axis mark the exact t–value of each study. In Panel b, field is defined by journal of publication.

income population (Milligan 2005; Riphahn and Wiynck 2017). This could be due to differential preferences and thresholds for "child quality," and/or to socioeconomic differences in the planning of births, and in what information is used in the planning process (Hayford and Guzzo 2016).

Selectivity and sources of bias

Publication bias and evaluation of minuscule changes: Opposing sources of bias. Publication bias, wherein studies rejecting the null are more likely to be published, means that "published evidence is unrepresentative of reality" (Ioannidis 2008; Simonshohn et al. 2014). Brodeur et al. (2020) find substantial evidence of p-hacking in the quasi-experimental literature, particularly in reporting of results from DiD designs. Figure 4a shows a heaping of *t*-values just above the thresholds for significance at the 5 and 10 percent level also in our selection of studies. In line with this tendency, an Egger test (Egger et al. 1997) rejected the null hypothesis, suggesting that publication bias may be present ($p < 0.001$). This means that studies that find no effect on fertility are more likely to remain unpublished. In short, publication bias contributes to the overestimation of the importance of policy for fertility.

Subsample estimation amplifies the challenges of multiple testing.[5] A number of the included studies test and report subsample results along a

number of stratification lines. In general, subsample differences that are significant are more likely than the nonsignificant results to be emphasized in abstracts and conclusions. To the best of our knowledge, none of the included studies had preregistered their analysis plan, and we cannot determine if the choice of stratification variables was post hoc motivated, and if the complete results of subsample tests are reported.

As studies that reject the null are likely easier to publish, one could assume that researchers investigate only large reforms that plausibly affect fertility, ignoring smaller changes. However, our review of studies suggests the opposite: often, the incremental policy change is so small, or the estimated reform effect captures such a small part of the total reform effect, that the estimated reform effect would likely be zero even when the true total reform effect is nonzero. When reforms large enough to be "critical junctures" (Neyer and Andersson 2008) are evaluated in analytical designs that capture a large share of the total reform effect, fertility responses are systematically larger.

In the wake of "economic imperialism": Does field of publication matter? Figure 4b shows the distribution of studies by topic and field of publication. All studies on cash-transfer reforms, well-suited to test predictions from the microeconomic theory of the family (Becker 1991), are published in economics journals. Childcare and parental leave as fertility determinants have generally attracted more interest among demographers and are published in demographic and economic journals. Studies of health-service reforms appear in journals of economics, health, and demography.

Judgment calls in data preparation and sample definition have substantial bearing on the results in quasi-experimental analysis (Huntington-Klein et al. 2021). Field may matter for these choices; for instance, economists have traditionally put relatively more emphasis on completed fertility, while demographers to a larger extent examine temporary ("tempo") effects. We do not find a tendency of such differences by field in our review. Data limitations rather than field of publication outlet set the time horizon. Studies across fields compare long- and short-term effects.

Conclusion

In this review, we have summarized studies on the effects of policies on fertility based on an extensive and systematic search of both published articles and working papers (>17,000 screened). In sum, our bird's-eye view of quasi-experimental results corroborates an observation long made by demographers: family-friendly policies *do* contribute to high fertility. Our overview suggests that differences in fertility across countries have emerged, at least partly, as a result of family policies. Although this hypothesis is by no means new, the synthesized causal evidence that underpins it is.

In 2007, Gauthier suggested that effects of policies were mainly temporary, but concluded that evidence was limited, and more solid causal evidence was needed (see also Thévenon and Gauthier 2011). Fourteen years later, we have synthesized results based on new developments in empirical causal modeling, concluding that policies may be more important for fertility than previously thought. This conclusion emerges only when we put most weight on the studies that capture changes in policies plausibly large enough to impact fertility, and, as in previous reviews, we conclude that the effects of cash transfers are likely transitory. When assessing whether effects impact completed fertility, we have considered multiple dimensions, including duration of measurement period, and whether effects occur at the end of the reproductive ages and/or at higher parities.

We note that quasi-experiments better than cross-country comparisons can pinpoint policy effects by zooming in on the affected population, avoiding the fact that groups for whom the policy change is irrelevant wash out the measurement of effects. As such, our review of the literature underlines that detailed microdata are needed to understand the social processes that link policy change to fertility (Neyer and Andersson 2008). However, policy *effects* are often different from the microlevel association between policy *uptake* and subsequent fertility. The association between paternity leave uptake and parity progression is positive (Duvander et al. 2010, 2019), while the causal effect is nil or negative. For childcare supply and fertility, the naïve association with fertility is negative, while the causal effect is positive (Rindfuss et al. 2010). Individuals select into the care arrangement that best suits their needs and preferences. To separate this selection from policy effects, careful statistical modeling is required.

Previous reviews have remained inconclusive on the effects of parental leave, summarizing them as "weak," "mixed," or "not unequivocal" (e.g., Gauthier 2007; Sleebos 2003; Sobotka et al. 2020). Contrary to this, we find lasting and substantial effects on fertility both for large reforms of public childcare *and* parental leave. These two policies go a long way in alleviating the conflict between women's roles as workers and mothers, and evidence of effects is particularly present from contexts where this conflict presumably was quite high (Norway in the 1970s and Central Europe in the 2000s). When implemented together, they allow parents to retain their jobs while they care for their newborn, and then go back to paid work when their child is ready to be cared for by others. As such, these policies may both increase fertility by reducing opportunity costs, and increase maternal labor supply (Gauthier 2007; Goldscheider et al. 2015). Very long compensated parental leave may, however, cement traditional gender roles and strengthen the "glass ceiling" faced by women in the labor market (Datta Gupta et al. 2008).

What about the effect of policies in pacing up the last step in the "gender revolution" (Goldscheider et al. 2015)? For paternity leave, a negative effect emerged in Spain, while no effect is found in the Nordic countries.

Unlike childcare and parental leave, effects of paternity quotas on mothers' opportunity costs are mainly indirect, contingent on fathers responding by permanently increasing their efforts at home. Furthermore, men may be reluctant to father (many) children if responsibilities at home increase. We emphasize that evidence is scarce, and that small incremental changes in maternity leave had no effect, either.

Women's increased education and employment have contributed to raising the age at first childbirth in several countries. For European countries struggling with low fertility, offering assisted reproductive treatment to women of all ages for which the success rate is of a meaningful size might be a policy option. As an increasing share of births happen in the late 30s, the potential for assisted reproduction to impact TFRs also increases (Sobotka et al. 2020). These births will be to parents who are strongly motivated to raise children. Presumably, this will have a positive impact on the development and adult functioning of the additional children born.

Family policies usually aim to balance several goals, and awareness of population heterogeneity in responses is important for understanding their total impact on society. It is striking that universal transfers seem to have the largest fertility effects among well-off families who need them the least. Earnings-related parental leaves imply a relatively large redistribution toward higher earning couples (Dahl et al. 2016), and this group also sees the most positive effects on fertility. In-kind services, in contrast, do not show the same social pattern of fertility effects. High-quality health services and childcare early in life have long-term positive effects on health, educational attainment, and earnings, particularly for children from poorer families (Campbell et al. 2014; De Haan and Leuven 2020). If reducing inequality is a goal, in-kind services such as child and health care are preferable. In practice, this means that cheap health services should be prioritized over cash transfers, and improvements in childcare coverage and quality chosen over compensating very long parental leaves.

Our review further suggests that the symbolic meaning and/or signaling effect of announcing pronatalist policies should not be underestimated. The largest fertility effects of cash transfers and parental leave expansions emerged where reforms have explicitly aimed to increase fertility (e.g., Milligan 2005; Raute 2019).[6] However, pronatalist-motivated reforms also tend to be larger, which may fully explain the larger effects. Given the deep controversies regarding pronatalism in population policies (Schultz 2015), attempts to amplify effects with a pronatalist message may also backfire (Botev 2015). Yet, to embed public support in their decision-making, individuals need to be informed about policies. In line with this, we see evidence of announcement effects, meaning that policies take effect when they are announced but not yet implemented (e.g., Mörk et al. 2013). The comparatively large fertility effects of field experiments may reflect the fact that they are more clearly announced to the public than other reforms

tend to be. Alternatively, rational individuals may act on information about changes in future costs of childbearing or decide to have a child now rather than later if it is temporarily cheaper because of a field experiment.

The post-2009 fertility fall has been steep even in the Nordic countries (Comolli et al. 2021), despite virtually unchanged family policies. At the same time, Germany has seen a marked fertility increase, as their family policies have expanded in the direction of the Nordic model. This could indicate a "hedonic treadmill" (Brickman and Campbell 1971), where individuals over time adapt expectations and take support for granted, so that effects wane. In the Nordic countries, today's parents-to-be grew up with extensive parental leave and accessible public childcare, and in their fertility choices take for granted this sense of security from public support, as well as almost uninterrupted careers for men and women (Ellingsæter and Pedersen 2016; Cools and Strøm 2020). A retraction of family policies would, therefore, likely cause a sense of relative deprivation and reduce fertility.

Countries with extensive family policies already have the best documented measures in place, though their fertility effects may be waning over time. Here, innovative and creative solutions for the future will not necessarily be directly evidence-based. With increasing educational attainment among both men and women all over the globe, and a reluctance to have children before gaining a foothold in the labor market, it seems difficult to avoid that both families and careers are established in the same "rush hour of life" (Zanella et al. 2019). Based on theory and survey data on preferences, one can think about measures that further ease the balance between work and care, such as additional improvements in childcare accessibility, opening hours, or quality. An improved digital infrastructure for remote work in the wake of the COVID-19 pandemic may also ease work-life compatibility in the long run for some parents.

We conclude our review with a caution and a plea to policymakers. We found clear indications of publication bias, as identified previously for a broad range of fields. Alone, this suggests that the literature exaggerates the effects of policy on fertility. However, many of the policies analyzed are too small to plausibly have measurable effects on fertility; failing to account for this could lead to the false conclusion that policies barely impact fertility. Our findings call for further inquiry into publication bias for (quasi-)experimental studies of fertility, but also for demographic studies based on other quantitative techniques in general. It also underlines the fact that detailed preregistration of planned analyses is crucial to the credibility of quasi-experimental literature in the future. Field experiments on policies are rarely carried out, as they are expensive and politically controversial, which means that evaluations typically have to rely on quasi-experimental study designs. If there is political interest in the consequences of a policy, policy change should be implemented in a fashion that facilitates a statistically strong evaluation in a quasi-experimental design.

DATA AVAILABILITY STATEMENT

The data set of effect sizes, with information about rescaling, is available at Figshare, https://doi.org/10.6084/m9.figshare.15128961.

NOTES

This work was supported by the Norwegian Ministry of Children and Families, the Ministry of Health and Care Services, the Ministry of Labor and Social Affairs, and the Ministry of Education and Research under the "Determinants of falling fertility" project; the Norwegian Research Council under the Centers of Excellence funding scheme (#262700). Bergsvik's work was also supported by Norwegian Research Council through the FAMGEN project (#236926); Fauske's work was also supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 818420). We are grateful to Martin E. Andresen, Anna Godøy, Torleif Halkjelsvik, Juho Härkönen, Øystein Kravdal, Trude Lappegård, Kelly Musick, Torkild Hovde Lyngstad, and Kenneth A. Wiik for valuable comments. We would also like to thank librarian Gyri Straumann at the Department of Reviews and Health Technology Assessments, Norwegian Institute of Public Health, for conducing the basic search, and for subsequent guidance on databases and searching tools.

1 Despite specific additional searches, no studies matching our selection criteria were found on housing.

2 Overviews targeted at demographers (Moffitt 2003; Bhrolcháin and Dyson 2007; Engelhardt et al. 2009) or social sciences in general (Angrist and Pischke 2008; Morgan and Winship 2015) provide more details on general applications.

3 Ang classifies earlier years as treatment years and does not provide a description of pre-trends. The pattern of effects by parity diverges markedly from what one should expect from the parity variation in transfer increase.

4 Treatment is defined by place of residence, and if individuals are allowed to move (as in Milligan 2005), the migration of more family-oriented women to regions with more extensive family policies may bias estimates upward. If women are locked into their pre-reform region, as in Galloway and Hart (2015), the measurement error in place of residence will increase over time, so that reform effects are more precisely estimated in the short term than in the long term. For studies comparing Quebec to the rest of Canada, an added complication is that a number of Quebec-specific family-policy reforms have been implemented over time. When Parent and Wang (2007) find an effect of family policies in 1970 in Quebec, this questions the assumptions underlying the analyses of Ang (2015) and Milligan (2005), namely, that trends in Quebec and Canada were parallel prior to reforms in the mid-1980s.

5 If tests are performed in 10 subgroups, every second study will have a false positive (Type I error) at the 5 percent significance level.

6 There is also supporting evidence that the anti-natalist message that accompanied welfare cutbacks reduced fertility in itself (Jagannathan et al. 2010).

References

Abramowitz, Joelle. 2018. "Planning Parenthood: The Affordable Care Act Young Adult Provision and Pathways to Fertility." *Journal of Population Economics* 31(4): 1097–1123.

Andersson, Gunnar, Marit Rønsen, Lisbeth B. Knudsen, Trude Lappegård, Gerda Neyer, Kari Skrede, Kathrin Teschner, and Andres Vikat. 2009. "Cohort Fertility Patterns in the Nordic Countries." *Demographic Research* 20(14): 313–351.

Ang, Xiaoling L. 2015. "The Effects of Cash Transfer Fertility Incentives and Parental Leave Benefits on Fertility and Labor Supply: Evidence from Two Natural Experiments." *Journal of Family and Economic Issues* 36(2): 263–288.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press. ISBN: 9780691120355.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con Out of Econometrics." *Journal of Economic Perspectives* 24(2): 3–30.

Apostolova-Mihaylova, Maria, and Aaron Yelowitz. 2018. "Health Insurance, Fertility, and the Wantedness of Pregnancies: Evidence from Massachusetts." *Contemporary Economic Policy* 36(1): 59–72.

Battistin, Erich, Michele De Nadai, and Mario Padula. 2015. "*Roadblocks on the Road to Grandma's House: Fertility Consequences of Delayed Retirement*." Working Papers 748, Queen Mary University of London, School of Economics and Finance.

Bauernschuster, Stefan, Timo Hener, and Helmut Rainer. 2016. "Children of a (Policy) Revolution: The Introduction of Universal Child Care and Its Effect on Fertility." *Journal of the European Economic Association* 14(4): 975–1005.

Becker, Gary S. 1991. *A Treatise on the Family*. Cambridge: Harvard University Press.

Bhrolcháin, Máire Ní, and Tim Dyson. 2007. "On Causation in Demography: Issues and Illustrations." *Population and Development Review* 33(1): 1–36.

Botev, Nikolai. 2015. "Could Pronatalist Policies Discourage Childbearing?" *Population and Development Review* 41(2): 301–314.

Brickman, Philip, and Donald T. Campbell. 1971. "Hedonic Relativism and Planning the Good Science." In *Adaption Level Theory: A Symposium*, edited by Mortimer H. Appley, 287–302. New York: Academic Press.

Brodeur, Abel, Nikolai Cook, and Anthony G. Heyes. 2020. "Methods Matter: P-Hacking and Causal Inference in Economics." *American Economic Review* 110(11): 3634–3660.

Buttner, Thomas, and Wolfgang Lutz. 1990. "Estimating Fertility Responses to Policy Measures in the German Democratic Republic." *Population and Development Review* 16(3): 539–555.

Campbell, Frances, Gabriella Conti, James J. Heckman, Seong H. Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science* 343(6178): 1478–1485.

Cannonier, Colin. 2014. "Does the Family and Medical Leave Act (FMLA) Increase Fertility Behavior?" *Journal of Labor Research* 35(2): 105–132.

Cherlin, Andrew J. 2016. "A Happy Ending to a Half-Century of Family Change?" *Population and Development Review* 42(1): 121–129.

Comolli, Chiara L., Gerda Neyer, Gunnar Andersson, Lars Dommermuth, Peter Fallesen, Marika Jalovaara, Ari K. Jónsson, Martin Kolk, and Trude Lappegård. 2021. "Beyond the Economic Gaze: Childbearing during and after Recessions in the Nordic Countries." *European Journal of Population* 37: 473–520.

Cools, Sara, Jon H. Fiva, and Lars J. Kirkebøen. 2015. "Causal Effects of Paternity Leave on Children and Parents." *Scandinavian Journal of Economics* 117(3): 801–828.

Cools, Sara, and Marte Strøm. 2020. "Ønsker om barn - en spørreundersøkelse om fertilitet, arbeidsliv og familiepolitikk. [Child wishes – a survey on fertility, working life and family policy.]" *Rapport 5*, Institutt for Samfunnsforskning, Oslo.

Crimmins, Eileen M. 1985. *The Fertility Revolution: A Supply-Demand Analysis*. Chicago, IL: University of Chicago Press.

Cygan-Rehm, Kamila. 2016. "Parental Leave Benefit and Differential Fertility Responses: Evidence from a German Reform." *Journal of Population Economics* 29(1): 73–103.

Dahl, Gordon B., Katrine V. Løken, Magne Mogstad, and Kari V. Salvanes. 2016. "What is the Case for Paid Maternity Leave?" *Review of Economics and Statistics* 98(4): 655–670.

Datta Gupta, Nabanita, Nina Smith, and Mette Verner. 2008. "The Impact of Nordic Countries' Family Friendly Policies on Employment, Wages and Children." *Review of Economics of the Household* 6(1): 65–89.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110(9): 2964–2996.

De Haan, Monique, and Edwin Leuven. 2020. "Head Start and the Distribution of Long-Term Education and Labor Market Outcomes." *Journal of Labor Economics* 38(3): 727–765.

Deleire, Thomas, Leonard M. Lopoo, and Kosali I. Simon. 2011. "Medicaid Expansions and Fertility in the United States." *Demography* 48(2): 725–747.

Duvander, Ann-Zofie, Trude Lappegård, and Gunnar Andersson. 2010. "Family Policy and Fertility: Fathers' and Mothers' Use of Parental Leave and Continued Childbearing in Norway and Sweden." *Journal of European Social Policy* 20(1): 45–57.

Duvander, Ann-Zofie, Trude Lappegård, and Mats Johansson. 2020. "Impact of a Reform Towards Shared Parental Leave on Continued Fertility in Norway and Sweden." *Population Research and Policy Review* 39: 1205–1229.

Duvander, Ann-Zofie, Trude Lappegård, Synøve N. Andersen, Ólöf Garðarsdóttir, Gerda Neyer, and Ida Viklund. 2019. "Parental Leave Policies and Continued Childbearing in Iceland, Norway, and Sweden." *Demographic Research* 40: 1501–1528.

Egger, Matthias, George D. Smith, Martin Schneider, and Christoph Minder. 1997. "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *BMJ* 315(7109): 629–634.

Ellingsæter, Anne L., and Eirin Pedersen. 2016. "Institutional Trust: Family Policy and Fertility in Norway." *Social Politics: International Studies in Gender, State & Society* 23(1): 119–141.

Engelhardt, Henriette, Hans-Peter Kohler, and Alexia Fürnkranz-Prskawetz, eds. 2009. *Causal Analysis in Population Studies*. Amsterdam, The Netherlands: Springer.

Esping-Andersen, Gøsta. 1990. "The Three Political Economies of the Welfare State." *International Journal of Sociology* 20(3): 92–123.

Esping-Andersen, Gøsta, and Francesco C. Billari. 2015. "Re-theorizing Family Demographics." *Population and Development Review* 41(1): 1–31.

Farré, Lídia, and Libertad González. 2019. "Does Paternity Leave Reduce Fertility?" *Journal of Public Economics* 172: 52–66.

Fauske, Agnes, Rannveig K. Hart, and Janna Bergsvik. 2020. Protocol for 'Effects of Policy Effects on Fertility: A Systematic Review of (Quasi)Experiments'. *SocArXiv*, January 18 10.31235/osf.io/t8vsg.

Ferrera, Maurizio. 1996. "The Southern Model of Welfare in Social Europe." *Journal of European Social Policy* 6(1): 17–37.

Galloway, Taryn, and Rannveig K. Hart. 2015. "*Effects of Income and the Cost of Children on Fertility. Quasi-Experimental Evidence from Norway*." Discussion Papers 828, Statistics Norway, Research Department.

Gathmann, Christina, and Björn Sass. 2018. "Taxing Childcare: Effects on Childcare Choices, Family Labor Supply, and Children." *Journal of Labor Economics* 36(3): 665–709.

Gauthier, Anne H. 2002. "Family Policies in Industrialized Countries: Is There Convergence?" *Population* 57(3): 447–474.

Gauthier, Anne H. 2007. "The Impact of Family Policies on Fertility in Industrialized Countries: A Review of the Literature." *Population Research and Policy Review* 26(3): 323–346.

Gauthier, Anne H., and Jan Hatzius. 1997. "Family Benefits and Fertility: An Econometric Analysis." *Population Studies* 51(3): 295–306.

Goldscheider, Frances, Eva Bernhardt, and Trude Lappegård. 2015. "The Gender Revolution: A Framework for Understanding Changing Family and Demographic Behavior." *Population and Development Review* 41(2): 207–239.

Goldthorpe, John H. 1998. "Rational Action Theory for Sociology." *The British Journal of Sociology* 49(2): 167–192.

González, Libertad. 2013. "The Effect of a Universal Child Benefit on Conceptions, Abortions, and Early Maternal Labor Supply." *American Economic Journal: Economic Policy* 5(3): 160–188.

Groves, Lincoln H., and Leonard M. Lopoo. 2018. "Federal Financial Aid and Family Formation: Examining the Social Security Student Benefit Program." *Journal of Family and Economic Issues* 39(3): 436–444.

Groves, Lincoln H., Sarah Hamersma, and Leonard M. Lopoo. 2018. "Pregnancy Medicaid Expansions and Fertility: Differentiating between the Intensive and Extensive Margins." *Population Research and Policy Review* 37(3): 461–484.

Greenland, Sander. 1990. "Randomization, Statistics, and Causal Inference." *Epidemiology* 1(6): 421–429.

Hakim, Catherine. 2003. "A New Approach to Explaining Fertility Patterns: Preference Theory." *Population and Development Review* 29(3): 349–374.

Hart, Rannveig K., Synøve N. Andersen, and Nina Drange. 2019b. "*Effects of Extended Paternity Leave on Union Stability and Fertility*." Discussion Papers 899, Statistics Norway, Research Department.

Hart, Rannveig K., Janna Bergsvik, and Agnes Fauske. 2019a. "*A Systematic Review of the Causal Effects of Public Policies on Birth Rates in Europe."*. Northern America and Australia. *PROSPERO 2019* CRD42019134807 www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019134807.

Hayford, Sarah R., and Karen B. Guzzo. 2016. "Fifty Years of Unintended Births: Education Gradients in Unintended Fertility in the US, 1960–2013." *Population and Developement Review* 42: 313–341.

Heim, Bradley, Ithai Lurie, and Kosali Simon. 2018. "The Impact of the Affordable Care Act Young Adult Provision on Childbearing: Evidence from Tax Data." *Demography* 55(4): 1233–1243.

Hellstrand, Julia, Jessica Nisén, and Mikko Myrskylä. 2020. "All-Time Low Period Fertility in Finland: Demographic Drivers, Tempo Effects, and Cohort Implications." *Population Studies* 74(3): 315–329.

Hilgeman, Christin, and Carter T. Butts. 2009. "Women's Employment and Fertility: A Welfare Regime Paradox." *Social Science Research* 38(1): 103–117.

Hoem, Jan M. 1990. "Social Policy and Recent Fertility Change in Sweden." *Population And Development Review* 16(4): 735–748.

Hoem, Jan M. 2008. "Overview Chapter 8: The Impact of Public Policies on European Fertility." *Demographic Research* 19: 249–260.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396): 945–960.

Hotz, Joseph V., Jacob A. Klerman, and Robert J. Willis. 1997. "Chapter 7 The Economics of Fertility in Developed Countries." In *Handbook of Population and Family Economics*, 275–347. 1 Part A, Elsevier.

Huntington-Klein, Nick, Andreu Arenas, Emily Beam, Marco Bertoni, Jeffrey R. Bloem, Pralhad Burli, Naibin Chen, Paul Grieco, Godwin Ekpe, Todd Pugatch, Martin Saavedra, and Yaniv Stopnitzky. 2021. "The Influence of Hidden Researcher Decisions in Applied Microeconomics." *Economic Inquiry* 59(3), 1–17. https://doi.org/10.1111/ecin.12992

Ioannidis, John P. A. 2008. "Why Most Discovered True Associations are Inflated." *Epidemiology* 19(5): 640–648.

Jagannathan, Radha, Michael J. Camasso, and Carol Harvey. 2010. "The Price Effects of Family Caps on Fertility Decisions of Poor Women." *Journal of Social Service Research* 36(4): 346–361.

Jalovaara, Marika, Gerda Neyer, Gunnar Andersson, Johan Dahlberg, Lars Dommermuth, Peter Fallesen, and Trude Lappegård. 2019. "Education, Gender, and Cohort Fertility in the Nordic Countries." *European Journal of Population* 35: 563–586.

Joyce, Theodore, Robert Kaestner, and Florence Kwan. 1998. "Is Medicaid Pronatalist? The Effect of Eligibility Expansions on Abortions and Births." *Family Planning Perspectives* 30(3): 108–113, 127.

Kalwij, Adriaan. 2010. "The Impact of Family Policy Expenditure on Fertility in Western Europe." *Demography* 47(2): 503–519.

Lalive, Rafael, and Josef Zweimüller. 2009. "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments." *Quarterly Journal of Economics* 124(3): 1363–1402.

Laughlin, Lynda. 2013. "*Who's Minding the Kids? Child Care Arrangements: Spring 2011*." Current Population Reports P70-135. Washington, DC: U.S. Census Bureau.

Leibowitz, Arleen. 1990. "The Response of Births to Changes in Health Care Costs." *Journal of Human Resources* 25(4): 697–711.

Leitner, Sigrid, Ilona Ostner, and Cristoph Scmitt. 2008. "Family Policies in Germany." In *Family Policies in the Context of Family Change: The Nordic Countries in Comparative Perspective*, edited by Ilona Ostner and Cristoph Scmitt, 175–202. Wiesbaden: VS Verlag für Sozialwissenschaften

Liu, Qian, and Oskar N. Skans. 2010. "The Duration of Paid Parental Leave and Children's Scholastic Performance." *The B E Journal of Economic Analysis & Policy* 10(1): Article 3.

Machado, Matilde P., and Anna Sanz-de-Galdeano. 2015. "Coverage of Infertility Treatment and Fertility Outcomes." *SERIEs* 6: 407–439.

Margolis, Rachel, Youjin Choi, Anders Holm and Nirav Mehta. 2021. "The Effect of Expanded Parental Benefits on Union Dissolution." *Journal of Marriage and Family* 83(1): 191–208.

McDonald, Peter. 2006. "Low Fertility and the State: The Efficacy of Policy." *Population and Development Review* 32(3): 485–510.

McLanahan, Sara. 2004. "Diverging Destinies: How Children are Faring under the Second Demographic Transition." *Demography* 41(4): 607–627.

Milligan, Kevin. 2005. "Subsidizing the Stork: New Evidence on Tax Incentives and Fertility." *Review of Economics and Statistics* 87(3): 539–555.

Moffitt, Robert. 2003. "Causal Analysis in Population Research: An Economist's Perspective." *Population and Development Review* 29(3): 448–58.

Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and Causal Inference*. Cambridge: Cambridge University Press.

Mörk, Eva, Anna Sjögren, and Helena Svaleryd. 2013. "Childcare Costs and the Demand for Children—Evidence from a Nationwide Reform." *Journal of Population Economics* 26(1): 33–65.

Neyer, Gerda, and Gunnar Andersson. 2008. "Consequences of Family Policies on Childbearing Behavior: Effects or Artifacts?" *Population and Development Review* 34(4): 699–724.

Neyer, Gerda, Arianna Caporali, and Nora S. Gassen. 2017. "*EU-Policies and Fertility: The Emergence of Fertility-Related Family Policies at the Supra-National Level*." Stockholm Research Reports in Demography 10, Stockholm University, Demography Unit.

Parent, Daniel, and Ling Wang. 2007. "Tax Incentives and Fertility in Canada: Quantum vs Tempo Effects." *Canadian Journal of Economics* 40(2): 371–400.

Presser, Harriet B., and Wendy Baldwin. 1980. "Child Care as a Constraint on Employment: Prevalence, Correlates, and Bearing on the Work and Fertility Nexus." *American Journal of Sociology* 85(5): 1202–1213.

Raute, Anna. 2019. "Can Financial Incentives Reduce the Baby Gap? Evidence from a Reform in Maternity Leave Benefits." *Journal of Public Economics* 169: 203–222.

Rindfuss, Ronald R., David K. Guilkey, S. Philip Morgan, Øystein Kravdal, and Karen B. Guzzo. 2007. "Child Care Availability and First-Birth Timing in Norway." *Demography* 44(2): 345–372.

Rindfuss, Ronald R., David K. Guilkey, S. Philip Morgan, and Øystein Kravdal. 2010. "Child-Care Availability and Fertility in Norway." *Population and Development Review* 36(4): 725–748.

Riphahn, Regina T., and Frederik Wiynck. 2017. "Fertility Effects of Child Benefits." *Journal of Population Economics* 30(4): 1135–1184.

Ruhm, Christopher J. 2019. "Shackling the Identification Police?" *Southern Economic Journal* 85(4): 1016–26.

Rønsen, Marit, and Kari Skrede. 2010. "Can Public Policies Sustain Fertility in the Nordic Countries? Lessons from the Past and Questions for the Future." *Demographic Research* 22: 321–346.

Schmidt, Lucie. 2005. "Infertility Insurance Mandates and Fertility." *American Economic Review* 95(2): 204–208.

Schmidt, Lucie. 2007. "Effects of Infertility Insurance Mandates on Fertility." *Journal of Health Economics* 26(3): 431–446.

Schultz, Susanne. 2015. "Reproducing the Nation: The New German Population Policy and the Concept of Demographization." *Distinktion: Journal of Social Theory* 16(3): 337–361.

Simonshohn, Uri, Leif D. Nelson, and Joseph P. Simmons. 2014. "P-Curve: A Key to the File-Drawer." *Journal of Experimental Psychology: General* 143(2): 534.

Sobotka, Thomas, Anna Matysiak and Zusanna Brzozowska. 2020. "*Policy Responses to Low Fertility: How Effective are They?"* Working Paper No. 1, UNFPA, Technical Division, Population & Development Branch.

Sleebos, Joelle. 2003. "Low Fertility Rates in OECD Countries: Facts and Policy Responses." *OECD Labour Market and Social Policy Occasional Papers* 15. https://doi.org/10.1787/568477207883.

Thévenon, Olivier and Anne H. Gauthier. 2011. "Family Policies in Developed Countries: A 'Fertility-Booster' with Side-Effects." *Community, Work & Family* 14(2): 197–216.

Thomson, Elizabeth, Trude Lappegård, Marcia Carlson, Ann Evans and Edith Gray. 2014. "Childbearing across Partnerships in Australia, the United States, Norway and Sweden." *Demography* 51(2):485–508.

United Nations. 2018. *World Population Policies 2015*. New York, NY: United Nations, Department of Economic and Social Affairs.

Vignoli, Daniele, Raffaele Guetto, Giacomo Bazzani, Elena Pirani, and Alessandra Minello. 2020. "A Reflection on Economic Uncertainty and Fertility in Europe: The Narrative Framework." *Genus* 76(1): 1–27.

Wood, Jonas, and Karel Neels. 2019. "Local Childcare Availability and Dual-Earner Fertility: Variation in Childcare Coverage and Birth Hazards Over Place and Time." *European Journal of Population* 35: 913–937.

Wood, Jonas, Karel Neels, and Jorik Vergauwen. 2016. "Economic and Institutional Context and Second Births in Seven European Countries." *Population Research and Policy Review* 35(3): 305–325.

World Bank. 2020. "Fertility Rate, Total (Births per Woman)." Spain, Italy, United States. Accessed May 5, 2020. https://data.worldbank.org/indicator/SP.DYN.TFRT.IN

Zanella, Marina, Bernhard Hammer, Alexia Prskawetz, and Jože Sambt. 2019. "A Quantitative Assessment of the Rush Hour of Life in Austria, Italy and Slovenia." *European Journal of Population* 35(4): 751–776.

Zavodny, Madeline, and Marianne P. Bitler. 2010. "The Effect of Medicaid Eligibility Expansions on Fertility." *Social Science & Medicine* 71(5): 918–924.