Does cutting child benefits reduce fertility in larger families? Evidence from the UK's two-child limit*

Mary Reader[†] Jonathan Portes[‡] Ruth Patrick[§]

Last updated: March 31, 2022

Abstract

We study the impact of restricting child-related social assistance to the first two children in the family on the fertility of third and subsequent births. As of April 2017, all third and subsequent born children to low-income families in the UK did not receive means-tested child benefits, amounting to a reduction in income relative to the previous system of approximately 3000 GBP a year per child. We use administrative births microdata and household survey data to estimate the impact of the two-child limit on higher-order births with a triple differences approach, exploiting variation over date of birth, socio-economic status, and birth order. We find some evidence that the policy led to a small decline in higher-order fertility among low-income families. However, compared to earlier research in the UK and elsewhere, largely based on benefit increases, the impact is small. This may be due to informational barriers or to other economic and social constraints affecting low income families. Our results imply that the main impact of cuts to child benefits is not to reduce fertility but to withdraw income from low-income families, with potential implications for child poverty.

Keywords: fertility, family size, social assistance, welfare reform

JEL: J13, J18, H31, H53

*We thank Mike Brewer, Kitty Stewart, Kate Andersen, other members of the Benefit Changes and Larger Families research team and our advisory board, CASE Researchers' Workshop attendees, and members of the Social Policy Quantitative Reading Group for helpful discussion and comments on previous drafts. We acknowledge funding from the Nuffield Foundation from grant FR-23208.

†Corresponding author. Email: `m.reader@lse.ac.uk`. Centre for Analysis of Social Exclusion, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, United Kingdom

‡Department of Political Economy, King's College London, Strand, London, WC2R 2LS, United Kingdom

§Department of Social Policy, University of York, Heslington, York, YO10 5DD, United Kingdom

This work was produced using statistical data from the Office for National Statistics. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

This project was funded by the Nuffield Foundation. The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation.

Contents

List of Tables

List of Figures

1 Introduction

How do financial incentives and constraints affect household fertility decisions? The basic economic model of fertility, as developed by Becker (1960), treats children as analogous to consumer durables, with associated costs and benefits. The implication is that government policies, in particular the treatment of children in the tax and benefit system, will change these costs and benefits and therefore affect fertility decisions. However, even in this simple model, the magnitude of these effects will depend both on individual and household preferences and on other variables affecting family incomes; they are therefore likely to depend on the specific economic and social context. The complexity of the factors involved, and the relative rarity of substantial changes to financial incentives that can be used to cleanly identify the impacts, mean that empirical estimates remain limited.

In this paper we use administrative and survey data to examine the impact of a major change to the treatment of children in the UK benefit system: the introduction of the "two-child limit" for cash benefits to low-income families. This change meant that means-tested child benefits – typically of the order of 2,845 GBP per year per child – were no longer payable for children born on or after 6 April 2017 to households that already had two or more children. While a handful of countries restrict means-tested child benefits to the third or fourth child, this policy was the first attempt internationally to cap child benefits at the second child (Longman, Patrick, Reeves & Stewart forthcoming). The two-child limit therefore represents a rare but significant policy experiment, which enables us to identify how financial incentives impact on fertility decision-making.

The impact of the two-child limit on families' fertility decisions is of considerable interest, both from a policy and a research perspective. In contrast to other changes to the UK benefit system for low-income families, this change was justified, at least in part, by the argument that benefit payments to low-income families with children incentivised higher levels of fertility, and that removing them would have the opposite effect (Treasury & DWP 2015). Implicitly, at least, the objective of the policy was not just to reduce expenditure but also to reduce fertility among larger families on a low-income. It was argued that families on a low-income should have to make fertility decisions based on what they can afford, and in the absence of financial incentives to have further children.

By contrast, opponents of the change made both normative arguments – that it was inappropriate for the state to attempt to reduce fertility among low-income households by cutting benefits, and that if a certain level of support per child was appropriate, it should apply to all children – and positive ones, that the policy change was unlikely to have a large impact on fertility decisions. The main impact of the change, opponents argued, would therefore be to increase poverty among larger families. Establishing to what extent these narratives are consistent with the empirical evidence is useful in informing both evaluation of the policy and wider debate on the future of the welfare system. It is also relevant to wider debates about trends in fertility in advanced economies: the change came as overall fertility rates in the UK were (after a mini "baby boom" in the 2000s) falling steadily, with the UK's total fertility rate falling from 1.9 in 2011 to 1.65 in 2019.

The nature of the two-child limit, and the manner of its introduction, also makes it well suited to causal identification of the impact of financial incentives in the welfare system on fertility. It was a large change, introduced at a single point in time, and it affected some groups, while leaving others (those with fewer than two children, or not receiving benefits) entirely unaffected. In 2021, 308,520 households (including 1.1

million children) were affected by the policy (Department of Work and Pensions and Her Majesty's Revenue and Customs 2021). The policy only applied to third or subsequent children who were born on or after 6 April 2017, creating a plausibly exogenous source of treatment variation both by number of children and by child's date of birth. We leverage this variation to isolate the causal impact of the two-child limit on fertility using a difference-in-difference-in-differences (or triple differences) strategy, exploiting variation over time, over socio-economic status, and over birth order.

Our results suggest that the impact of the policy change on fertility was relatively small. The period before and after the policy change saw steady falls in the overall number of births in the UK. However, at the point of change, we do not find evidence of a large reduction in the relative number of births among those affected. Our preferred triple-difference estimate is of a decline of approximately 0.36 percentage points (approximately 5 percent) in the probability of low-income occupation women with two or more children having a 'higher-order' (third or subsequent) child after the policy. Of the births that took place, we also see a slight moderation of the pre-existing trend of increasing relative higher-order fertility rates among low-income groups. However, the effect size is again small.

On the face of it, these results are surprising, and out of line with the findings of most previous research in the UK and elsewhere; they challenge our understanding of how financial incentives affect household decision-making. It also undermines the implicit policy rationale for the change: rather than causing a major reduction in the number of children born into low-income families, the main impact will have been to increase the depth and incidence of child poverty.

The paper is structured as follows. We briefly describe the increase, and subsequent reduction, in the generosity of the welfare system in the UK over the last two decades, and the context for the introduction of the two-child limit. We then review the evidence on the impact of the structure and generosity of the benefit system on fertility. Our principal contribution is our empirical analysis, based on a version of a triple differences strategy which uses comprehensive administrative data on 3 million births in England and Wales both before and after the policy change. We then discuss possible interpretations and implications of our analysis, from both a research and policy perspective.

2 Background

Since the establishment of the modern welfare state after the Second World War, the UK has supported families with children through the tax and benefit system through a combination of universal (Family Allowance and Child Benefit) and means-tested support, initially for families with essentially no source of income at all. The latter was extended to low-income working families with the introduction of Family Income Supplement in the 1970s, made significantly more generous in the 2000s with the introduction of Working Families Tax Credit, which was modelled on the US Earned Income Tax Credit.

The objective of this expansion of benefits in the 2000s was both to reduce poverty among low-income families (the then government had set itself an objective of reducing and then eliminating child poverty, as measured by relative income) and to increase employment. Though there is no evidence to suggest that increased fertility was an objective of these policies, in practice Brewer, Ratcliffe & Smith (2012) found that those most likely to have been affected by the changes in financial incentives saw a differential increase in fertility, with an estimated rise of 1.2 to 3.6 percentage points in the annual probability of those most affected having a child, equating to an increase of approximately 10,000 to 35,000 in annual births. This was a period in which overall fertility rates in the UK were rising, with an increase in annual births of over 100,000 a year in the 2000s. In ONS (2009)'s descriptive analysis of these trends, they attributed about a third of the rise to the greater affordability of children for low-income households (including improved childcare provision as well as higher benefit levels), though this was not a causal estimate.

After the financial crisis and ensuing recession of 2008-09, and the election of a Conservative-Liberal Democrat coalition in 2010, the government enacted significant cuts to welfare benefits as part of its wider programme of fiscal consolidation. The initial programme of cuts included changes to benefits for sick and disabled people, capping the amount of benefits that could be received by individual households, reductions in the uprating of benefits, the so-called 'bedroom tax' that restricted housing benefit payments to those with 'extra' bedrooms, and other measures. It did not focus explicitly on larger families, though overall larger families were more adversely affected, especially by the benefit cap (Stewart, Reeves & Patrick 2021, Gaffney 2015, Reed & Portes 2015).

While presented primarily as an economic necessity – the prevailing levels of benefits were described as unaffordable, given the impact of the recession on government borrowing – a parallel narrative also suggested that the benefit system had undesirable moral hazard effects. The Prime Minister, David Cameron, argued in 2011:

> "The benefit system has created a benefit culture. It doesn't just allow people to act irresponsibly, but often actively encourages them to do so." (UK Government 2021).

In conjunction with some sections of the press, the government sought to justify benefit cuts by arguing that very significant payments were being made to people who were either making dubious or fraudulent claims or who were (legally) "milking the system" (Gaffney & Portes 2013). Frequently, the latter referred to larger families or so-called "benefits broods"; such a case, and the wider background, is described at length in Jensen & Tyler (2015).

2.1 The two-child limit

After winning an outright majority in the summer of 2015, in July the Conservative government announced a further programme of benefit cuts, including the introduction of a "two-child limit"; while the existing system of tax credits (which was in the process of being replaced with a new, integrated benefit, Universal Credit) paid an equal amount for each child, in future payments would only be made in respect of the first two children for each household. That is, for children born on or after 6 April 2017, no child element would be paid (with a limited number of exceptions) if the household already had two or more children. Households are exempt from the two-child limit if their third child is a multiple birth, an adopted child, in a non-parental caring arrangement, or if they are the result of non-consensual conception. See Treasury & DWP (2015) for a more detailed description of the scope and impact of the policy.

While this policy change was part of a wider package of cuts with the overall aim of reducing expenditure, it was also explicitly justified by reference to the broader arguments above. The government described the objective of the policy as being "to ensure that families in receipt of benefits faced the same financial choices about having children as those supporting themselves solely in work" (Treasury 2015). Implicit in this was a view that these incentives did in fact drive fertility decisions. This view was made explicit in published Impact Assessments, which (based in part on Brewer, Ratcliffe & Smith (2012)) suggested that the policy change would result in reduced fertility, although no quantified estimates were made:

> "In practice people may respond to the incentives that this policy provides and may have fewer children. There is no evidence currently available on the strength of these effects although the Institute for Fiscal Studies found a relationship between support for children in the benefit system and childbearing...Given that families are aware of the policy they may make the choice not to have (further) children." (Treasury & DWP 2015)

Yet in response to a Work and Pensions Select Committee report on the two-child limit, the government appeared to deny that curbing fertility was a key policy aim:

> "This policy does not attempt to limit the number of children people have. Claimants are able to have as many children as they choose, in the knowledge of the support available." (Work and Pensions Select Committee 2019).

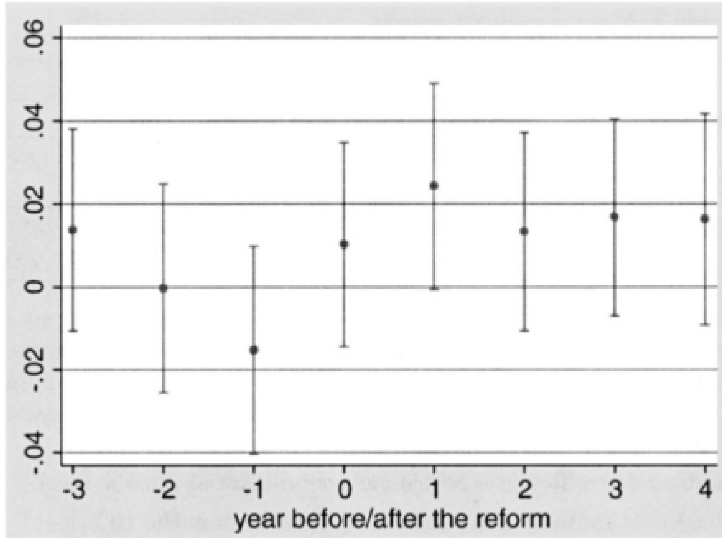
There were thus some inconsistencies in how the government defended the policy, with ambiguity over how far fertility changes were themselves an explicit policy objective, perhaps reflecting the political sensitivity around the intervention. But ultimately, irrespective of whether the reduction of fertility was intended or not, it is plausible that it may have achieved this in practice, given existing evidence on the link between welfare and fertility.

2.2 Existing evidence on welfare and fertility

As noted above, Brewer, Ratcliffe & Smith (2012) examined the impact of increases in the generosity of the UK benefit system in 1999. They used a difference-in-differences approach, exploiting the fact that the reforms were targeted at low-income households, and that the changes were likely to differentially affect the fertility of women in couples and single women because of the opportunity cost effects of the welfare-to-work element. Using data from the Family Resources Survey, they found no increase in births among single women, but an increase in births among coupled women on relatively low incomes, although, given the relatively small number of births observed in the data, there is considerable uncertainty around this estimate.

The key results are shown graphically in their Figure 2, reproduced below: there is an observable upward shift coinciding with the increase in the generosity of the system. It seems reasonable to conclude that there was indeed an impact, but the confidence intervals are large relative to its size and since the analysis looks at annual data, there is scope for other exogenous changes to affect the estimates. So while the central estimate is that the impact was an increase in the probability of having a child of approximately 2 percentage points per year in the affected groups (an increase in fertility rates of about 15 percent), a very high degree of uncertainty must be attached to this.

Fig. 2 Birth differentials between treatment and comparison groups



Source: Brewer, Ratcliffe & Smith (2012).

Meanwhile, Laroque & Salanié (2014) take a structural approach and, using data from France, find a substantial response of fertility to financial incentives in the French tax and benefit system. The response differs according to birth order, with first and (especially) third births particularly responsive. They use their estimates to simulate the impacts of an unconditional child credit of 150 Euros a month per child; they estimate that this would raise birth rates among the most affected group (i.e., those with two children) by 4.7 percentage points. These studies suggest quite large positive "elasticities" of fertility with respect to benefits. A number of other empirical studies in developed countries suggest positive fertility effects from the expansion of benefits (González & Trommlerová 2021). Evidence of negative fertility effects from benefit cuts or withdrawals is thinner, but there are some indications of this (Cohen, Dehejia & Romanov 2013, González & Trommlerová 2021).

These results are sensitive to institutional context. Research in the US is less suggestive of large impacts. Following the introduction of Temporary Assistance for Needy Families (TANF) programme in the 1990s, a number of states introduced measures to deny additional social assistance to any children born after a benefit claim was made. Research on these family caps has broadly found no effect on fertility (Joyce, Kaestner, Korenman & Henshaw 2004, Camasso & Jagannathan 2016, Kearney 2004, Dyer & Fairlie 2004). However, even leaving aside the very different social and economic context of the US, it is not clear that the two-child limit will have the same fertility effects as family caps. While the policies are similar in their aims, they are rather different in their application: for family caps, the determinant of whether you are capped is the timing of pregnancy relative to your benefits claim; for the two-child limit, it is the size of your family. The interaction with other aspects of the welfare system (notably work requirements and childcare) is also very different.

Fertility responses are also heterogeneous across the income distribution. The existence and magnitude of effects, therefore, depends on the specific targeting of the policy in question. Milligan (2005), using a triple differences approach to examine the introduction and subsequent withdrawal of a universal C$8,000 cash payment to families with newborn children in Quebec, found large fertility effects, with an extra C$1,000 leading to an increase of about 17 percent in fertility. However, the effects were heterogeneous in income, with considerably larger fertility responses among better-off families. While this may seem counterintuitive, it is in fact consistent with the canonical Becker model. Where families can choose to spend on both the quantity and the "quality" of children (where "quality" is defined as per-child spending), the impact of a simple per-child benefit is ambiguous. While the substitution effect (having children becomes less costly relative to working) unambiguously increases fertility, the income effect might either increase or decrease fertility, depending on whether the income elasticity of fertility is negative or positive (just as a cut in income taxes can either increase or decrease hours worked, depending on the income elasticity of demand for leisure). So while intuitively one would expect a cut in child benefits to reduce fertility, this need not be the case for all families (Milligan 2005, Cohen, Dehejia & Romanov 2013, Riphahn & Wiynck 2017).

In particular, if low-income families face a particular sharp "quality-quantity" trade off, they may respond to the higher incomes that result from a per-child benefit by reducing the number of children. That is, the increased income resulting from a per-child benefit for poorer families with one or two children may mean that they choose not to have a subsequent child.

However, for the policy change we examine – the removal of benefits for third and subsequent children – the income effect is less relevant. There is no income effect for the first two children, so the choice to have a third child should be driven solely by the substitution effect, and the impact of the policy change should be unequivocally negative.[1]

Our prior, therefore, based on previous research, and the nature of the policy change, is that the two-child limit is likely to have reduced fertility among affected households. There is no reason to expect the fertility elasticities with respect to financial incentives to be the same across time and place, but if the impacts were of similar magnitudes to those found in the papers we focus on in the discussion above, we would expect to see very large changes among affected households. The financial impact of the two-child limit is almost twice as large in PPP terms as that of the hypothetical benefit increase modelled by Laroque & Salanié (2014) (OECD 2021), implying an impact of more than 7 percentage points, and even larger compared to Milligan (2005). Similarly, the implied impact would be considerably larger than that suggested by the Brewer, Ratcliffe & Smith (2012) estimate: the two-child limit costs affected households by up to 55 GBP per week, considerably larger (in real terms) than the differential impacts of the 1999 changes. With this in mind, we turn to our empirical analysis.

3 Data

Our main data source is administrative birth registrations microdata from 2015-2019 in England and Wales. This provides individual-level data for all births (both live and stillbirths) in England and Wales over this period, with variables on the date of birth of the child, the number of previous live births born to the mother, multiple birth status, stillbirth status, maternal and paternal age, country of birth and postcode of residence of the mother. In addition, for a random 10 percent sample of live births, the data includes socio-economic status (SES), as defined by the National Statistics Socio-economic Classification (NS-SEC)'s measure of occupation (Office for National Statistics 2016). We create a measure of combined occupational

[1] In principle, the removal of benefits for the third child could, via the income effect, lead to some families choosing to have a fourth child. However, this is highly unlikely to be relevant over the time period we examine.

status by selecting the highest-ranked occupation of the mother or father, where available. Though we also use mother's occupation as a robustness check, family occupation is our preferred measure as it best captures the likelihood of household entitlement to means-tested benefits (to see the relationship between family occupation, the number of children and benefits receipt, see Figure 1 of the Supplementary Material). Among women aged 16-45 with two children, for example, 65 percent of those in routine occupations receive one or more of these benefits, relative to just 6 percent of those in higher professional and managerial occupations.

In the births data we also construct geographical variables to proxy the probability that a household will be affected by the two-child limit. We match each child's postcode of residence to its Lower Super Output Area (LSOA) – a small geographical area of approximately 1500 residents – and in turn match its LSOA to an Index of Income Deprivation score, a government measure which uses administrative benefits data to capture the proportion of the population in an area who are receiving means-tested benefits (Ministry of Housing, Communities and Local Government 2020).[2,3] We also match the residence of the child to published data at local authority level on the number of households who are affected by the two-child limit in 2021 (Department of Work and Pensions and Her Majesty's Revenue and Customs 2021). In combination with published data from the Annual Population Survey on the number of households by local authority as of 2019, we construct an local incidence variable that measures the percentage of total households in a local authority who are affected by the two-child limit (Office for National Statistics 2021).

Our second data source is the Annual Population Survey (APS). The APS is the largest household survey in the UK, with its annual sample including approximately 320,000 individual respondents in 122,000 households. The APS contains data at the individual level on the sex, age and NS-SEC occupation of the respondent, which benefits they receive if any, the number of dependent children under 16 in the family unit, and the type of family unit (e.g. single parent or couple family). The family unit is the relevant unit of observation for benefits receipt in the APS. We match respondents by family unit and construct a combined occupation measure that mirrors that of the births data, selecting the highest-ranked occupation of members of the family. We restrict our APS sample to women of approximate childbearing age (16-45) in England and Wales, the latter of which to be consistent with the geographical coverage of the administrative births data.

The administrative births data enable us to accurately capture the number of births at population level, with very low measurement and sampling error. However, the unit of observation in the births data is the child rather than the woman or couple deciding whether to have a child (our main population of interest). Women who do not have a child are, by definition, not captured in the births data. Additionally, the births data does not contain individual-level data on benefits receipt. We therefore use the APS for two purposes: first, to calculate the total number of women in different population subgroups (in receipt of benefit, or not, and by number of children), so that we can compute average fertility rates for those groups; and second,

[2] We use the joint Income Deprivation Index from 2019 which covers both England and Wales.

[3] These benefits are very similar to, and in fact exhaustively cover, the benefits through which the two-child limit is implemented. This includes those in receipt of Income Support, income-based Jobseeker's Allowance, income-based Employment and Support Allowance, Pension Credit (Guarantee); Universal Credit where no adult is classed within the 'Working – no requirements' conditionality group; those not already covered in receipt of Working Tax Credit or Child Tax Credit with an equivalised income (excluding housing benefit) below 60 per cent of the national median before housing costs; and asylum seekers in receipt of subsistence support, accommodation support, or both (Ministry of Housing, Communities and Local Government 2020).

to quantify the relationship between occupation and benefits receipt, so that we can use occupation as a robust proxy for benefits receipt in our analysis. We are unfortunately not able to link the two datasets together; instead, we draw on them separately to estimate the relevant numerators and denominators for the probability of having a child, first by benefits receipt and family size, and then by occupation and family size, in order to conduct a causal triple differences analysis.

We focus on the years 2015-2019 in our analysis. This gives an approximately equal amount of time either side of the introduction of the policy. The recording of benefits receipt in the APS changed considerably in 2015, so it is difficult to construct a consistent series prior to that. We do not use 2020 data on account of the possible impact of Covid on fertility, which may introduce systematic bias into our results by affecting low-income and larger families more (Reader & Andersen 2022).

In Table 1 we present descriptive statistics for the administrative births data for the universe of all births in England and Wales from 2015 to 2019 (N= 3,380,560). 15.7 percent of births are to single mothers and 3 percent are multiple births (i.e., twins or triplets). Mean maternal and paternal age is 30.5 and 33.4 years respectively. The mean income deprivation score is 15.0, which means that on average 15 percent of the population in a given neighbourhood receive means-tested benefits. In Table 2 we include a breakdown of the number of women in our population of interest – women aged 16-45 in England and Wales – using weighted annual population estimates from the Annual Population Survey (unweighted N=265,070). There are approximately 11.2 million women in our population of interest, 1.1 million (on average across our chosen time period) of whom have two or more children and are in receipt of the benefits affected by the two-child limit.[4] Approximately 1.6 million women have two or more children and are in what we define below as low-income (or high benefit receipt) occupations.

It is worth noting three key features of the aggregate data. First, the number of births by year has been decreasing throughout the period covered by the data; second, the proportion of births that are 'higher-order' (that is, third or subsequent live-born children to the same mother) has also fallen. Figure 1 charts the number of total births (left axis) alongside the number of higher-order births (right axis) in England and Wales by year from 2013 to 2019. Total births have fallen especially sharply since 2016; meanwhile higher-order births have been falling steadily and consistently since 2013. Finally, fertility is higher among low-income groups (see Supplementary Material Table A1).

4 Method

Our objective is to identify the impact of the two-child limit on fertility of those 'treated' by the policy – that is, women (and/or couples) who, if they were to have a child after the implementation date of the policy, were no longer eligible for child-related benefits for that child. Since the policy applied only to those in receipt of benefits with two or more existing children and was rolled out over time, a difference-in-differences setup is a natural choice for an identification strategy. The policy does feature exogenous variation in treatment status by child's date of birth, which makes a regression discontinuity design a possibility. However, once we restrict to third and subsequent births and low-income families, the sample sizes involved do not give

[4]These benefits are Universal Credit, tax credits, housing benefit, Jobseeker's Allowance (JSA), and income support.

sufficient power for this approach at small bandwidths.

Estimating a difference-in-differences approach requires us to define treatment and control groups, and to establish that the parallel trends assumption is satisfied (that is, that in the absence of treatment, the treatment and control group would have evolved similarly). The obvious candidate for the treatment group is low-income women (single parents or members of a couple household) with at least two children. It is less clear what the control group is. One candidate is low-income women with no children or one child, who, even after the policy change, could have a child and still receive an additional child-related benefit payment. But the parallel trends assumption seems implausible: there are numerous factors other than the policy change (other changes to the benefit system and wider social and demographic trends) that would be likely to change the desired number of children, and hence relative birth rates to women with different numbers of children.

An alternative candidate for the control group is high-income women with at least two children. Broader social trends on family size would arguably affect both these treatment and control groups in similar ways. However, again, it is plausible that trends in fertility may differ between low-income and high-income women.

Our preferred approach is therefore a difference-in-difference-in-differences (or triple differences) approach, which creates four groups, only one of which is treated, and which combines the two differences above. That is, the impact of the policy change is estimated by looking first at the change in fertility between low-income women with larger families and high-income women with larger families; and then at the change in fertility between low-income women with smaller families and high-income women with smaller families, and comparing these two changes.[5]

In principle, this should abstract both from differential changes in fertility between those on a low income and those not, and differential changes in fertility between larger and smaller families. This triple difference approach (which is possible because of the nature of the policy change, affecting only women with larger families on benefits, rather than just those on benefits or just those with large families) should enable us to be more confident that the parallel trends assumption is satisfied and should be more robust than a simple difference-in-differences strategy (Milligan 2005).

We implement the triple differences with two different approaches. First, we estimate differential changes in the probability of low income women having a higher-order (3+) child after the reform. Drawing on both APS and administrative births data, this approach enables us to focus on our main outcome of interest - the probability of women having a child - but may be affected by sampling and measurement error of the APS. In our second approach, we estimate differential changes in the probability of a child, having been born, being a higher-order (3+) birth to a low-income family after the reform. This approach draws exclusively on the administrative births data, but in doing so only enables us to observe the numerators of these probabilities - births.

[5] Note that it does not matter which difference is taken first; that is, we could equally compare fertility rates for low-income women with no or one child with those with two or more children; do the same for high-income women; and then compare the two differences. As the algebraic presentation below makes clear, the result is identical.

4.1 Differential changes in the probability of low-income women having a higher-order (3+) child

Our first approach is relatively standard: we compare changes in the probability of a low-income/high-income woman having a 3+/2- child before and after the reform. There are two main methodological choices here: how to define being on a low-income, and how to calculate the probability of having a child.

4.1.1 Defining low income

One of the three elements of variation in the triple differences is whether a birth is to a low-income family (and therefore a family likely to be receiving means-tested benefits) or not. Deciding how to measure this involves a trade-off between accuracy and bias. One obvious option is to use self-reported benefits receipt[6], since it should directly capture whether a family falls within the scope of the benefits system and therefore the two-child limit. This is the approach we take in Section 5.1.

However, benefits receipt may generate bias in our estimates over time for two reasons. First, benefits receipt is not exogenous either to having an additional child or to the policy itself. Most obviously, single women with no children are likely to be in work, and unlikely to qualify for means-tested benefits; but most single women with children do qualify, and in general the likelihood of claiming benefits rises with the number of children. [7] Moreover, the policy change itself impacts benefit entitlement, so some women with two children not on benefits who go on to have a third child would, prior to the policy change, qualify for benefits, but no longer do so. This subgroup should in principle be in the treatment group (since they are affected by the policy change) but the analysis above will not identify them as such.

Second, the characteristics of women on benefits has changed over the period in question due to the decline in the number of women claiming benefits over this period (DWP 2021, HMRC 2021). While this fall was primarily exogenous – it was not driven by the two-child limit, but rather by other reductions in benefit generosity, and rising employment – it was sufficiently large to affect the demographic composition of the different groups. While we have no reason to believe that this would bias our results, particularly as the triple differences exploits variation by birth order as well as benefits receipt, it is difficult to rule out the possibility entirely. This also makes interpretation harder, since the impact on families' decision making presumably depends on whether they expect to be affected by the policy change, which may in turn be determined by benefit history and expectations as well as current benefit status.

In Section 5.2 we therefore use a more exogenous, and less time-varying, measure of low-income: whether parents in a family are working in a low-income occupation or not. This is a fuzzy but valid instrument for the probability of being on benefits, as shown in Figure 1 of the Supplementary Material. Family occupation is strongly correlated with benefits receipt, but it is much less likely to change from year to year and (particularly at the family level) is less likely to be endogenous to the number of children or to the

[6]These benefits are the main means-tested benefits that are affected by the two-child limit: Universal Credit, tax credits, income support, and jobseekers' allowance. Note that we do not include child benefit, since it is not affected by the two-child limit.

[7]This is likely to explain at least in part the much higher 'risk' shown in the third column of Table 3; the denominator is an underestimate of the true size of the treatment group here.

policy change. It may therefore be better correlated with a woman's perception of the likelihood of being affected by the policy than benefits receipt at a point in time. Measuring treatment by reference to this more persistent indicator may give a better measure of "effective" treatment than benefits receipt at a point in time (just as SES may in some contexts be a better indicator of, say, living standards than contemporaneous income, if income is highly variable). Based on the observed probabilities in Figure 1 of the Supplementary Material, we group observations into those with a family occupation with a higher probability of benefits receipt (NS-SEC 3, 4, 5, 6, and 7) and those with a lower probability of benefits receipt (those in NS-SEC 1 and 2).

4.1.2 Estimating the probability of having a child

In this approach, our relevant outcome of interest is the probability of a woman or couple having a child. To estimate these probabilities, we need to separately identify the numerators and denominators for each group at each time point.

The denominator is relatively simple: it is the number of women who could potentially have children, grouped by whether they already have two or more children or not, and by whether they are on a low income (as defined by benefits receipt or family occupation). To construct the denominators we use the Annual Population Survey to construct annual weighted population estimates of the number of women aged 16-45 in the following categories: low-income women with 0-1 child; high-income women with 0-1 child; low-income women with two or more children; and high-income women with two or more children. This provides our estimate of the number of women in the four distinct groups (one treated; three non-treated) defined above.

We now need to estimate the numerator: the number of births in each group during each time period. Here we use the administrative births microdata to estimate the total number of births by year, family occupation, and the number of previous live births to the mother. This is preferable to relying on the Annual Population Survey alone; the administrative data (with family occupation) captures 10 percent of all births directly, much more than the APS and with greater reliability. In Section 5.1, where we use benefits receipt as a measure of treatment status, we need an estimate of the number of births by benefits receipt: this is complicated by the fact that the administrative births data does not provide any information on benefits eligibility. For this part of our analysis, we therefore use family occupation and the number of previous live births to estimate the probability that a given birth is to a family on benefits using the APS.[8] We use these probabilities to estimate the relevant numerators: the number of births by benefits receipt, higher-order (3+) status, and year. In Section 5.2, where we use low-income family occupation as a measure of treatment status, we simply count the number of births by family occupation, higher-order (3+) status, and year, to construct the relevant numerators.

For each group, we combine these numerators and denominators to calculate year-specific probabilities that a woman has a child. We then apply the triple differences approach to these probabilities to estimate whether fertility decreases differentially after the two-child limit among women on benefits with two or more existing children.

[8] Note that the APS does not include an exact variable for the number of previous live births to the mother, but the number of dependent children under 16 in the family is a relatively good proxy for this.

Algebraically, our empirical specification can be represented as follows:

$$P_{it} = \beta_0 + \beta_1 B_{it} + \beta_2 L_{it} + \beta_3 (B_{it} \times L_{it}) + \epsilon_{it}, \tag{1}$$

where P_{it} is the probability of individual i having a child in time period t, B_{it} is a dummy variable equal to one if the individual is on a low income (as defined either by benefits receipt or by low-income occupation), L_{it} is a dummy variable equal to one if the individual has two or more children, and ϵ_{it} is the error term. Our coefficient of interest is β_3: it represents the marginal impact of being on a low income and having two or more children on the probability of having a child.

Four groups can be defined by whether they are low-income and whether they have two or more children. The probability of having a child for each of these groups can be expressed as:

$$A : P_{it}|(B_{it}(1) \times L_{it}(1)) = \beta_0 + \beta_1 B_{it} + \beta_2 L_{it} + \beta_3 (B_{it} \times L_{it}) \tag{2}$$

$$B : P_{it}|(B_{it}(1) \times L_{it}(0)) = \beta_{0it} + \beta_1 B_{it} \tag{3}$$

$$C : P_{it}|(B_{it}(0) \times L_{it}(1)) = \beta_{0it} + \beta_2 L_{it} \tag{4}$$

$$D : P_{it}|(B_{it}(0) \times L_{it}(0)) = \beta_{0it} \tag{5}$$

We assume that β_0, β_1, β_2, and β_3 vary over time but remain constant between different groups (the "parallel trends") assumption. Accordingly, we denote the change in β_3 from time period $Post$ to period Pre as the following:

$$\beta_3{}' = \beta_{3post} - \beta_{3pre} \tag{6}$$

The difference in the probability of having a child between the $Post$ and Pre period is therefore the following for each group:

$$A : \beta_0{}' + \beta_1{}' + \beta_2{}' + \beta_3{}' \tag{7}$$

$$B : \beta_0{}' + \beta_1{}' \tag{8}$$

$$C : \beta_0{}' + \beta_2{}' \tag{9}$$

$$D : \beta_0{}' \tag{10}$$

We then have that $\beta_3{}'$ – the average impact of the change in benefits entitlement for a third or subsequent child, or the average treatment effect on the treated (ATT) – can be estimated by calculating $(A-B)-(C-D)$.

4.2 Differential changes in the probability of a child being a higher-order (3+) birth to a low-income family

The above approach, irrespective of whether we use benefits receipt or low-income occupation, relies on household survey data to construct the denominators in our fertility estimates. Since the number of women affected by the two-child limit since its introduction in 2017 is relatively small, it is possible that sampling or measurement error in the survey data attenuates our detection of fertility effects. We therefore undertake an alternative analysis in Section 5.3 which does not rely on survey data at all, but relies entirely on rich administrative data on the universe of all births in England and Wales. This has the advantage of allowing us to observe the entire population of births (or 10 percent of all births when we use occupational status), use both individual-level occupation data and local area geographical data on deprivation, and leverage the very large sample size of the births data. It also means that rather than having to use annual data, as above, we can use actual date of birth, which is relevant given the arbitrary date of birth cut-off for the two-child limit (6 April 2017).

In the administrative births data the universe is births, not women. Measuring the impact of the two-child limit therefore requires a different approach. We focus now not on the probability of a child being born to a woman in the treatment or control groups, but on the probability that a child, having been born, was born to a woman in the treatment or control groups, defined by reference to family occupation as above. The implicit assumption here is that, over the period we are looking at, changes in these probabilities will be driven by differential changes in fertility, rather than by changes in population demographics. The APS data suggests that this assumption is broadly valid: the number of women in the 16-45 population by group was stable over the period 2015-2019.

Again, this is a triple differences approach: we are comparing the change in the probability of an individual birth being higher-order between the treatment and the control groups; the probability of an individual birth being higher-order itself reflects the difference between the fertility rates of households with no or one child and those with two or more. Our empirical specification is the following:

$$L_{it} = \beta_0 + \beta_1 B_i + \beta_2 P_t + \beta_3 (B_i \times P_t) + \epsilon_{it} \tag{11}$$

, where L_{it} is a dummy variable equal to one if the child is the mother's third or subsequent birth and zero if the child is the mother's first or second birth; B_i is equal to one if the parents' combined occupational status is NS-SEC 3-7, and zero if it is NS-SEC 1-2; P_t is a dummy variable equal to one if the child is born on or after 6 April 2017 and zero otherwise; and ϵ_{it} is the error term.

The regression above therefore allows the probability of a higher-order birth to be influenced by the treatment variable – a proxy for the likelihood that the household is impacted by the two-child limit – and for that impact to vary before and after the treatment date. If the two-child limit is indeed having a differential negative impact on fertility, we would expect β_3 to be negative.

We also create two alternative treatment variables, using place of residence rather than family occupation as a proxy for benefits receipt. We do this, as described above, by matching in geographical data on the Index of Income Deprivation and the local incidence of the two-child limit. These variables are not dummy variables; they are continuous variables, where an increase in the variable implies a (broadly linear) increase in the probability that the mother is receiving benefits or, in the case of the "local incidence" variable, is impacted by the two child limit.

5 Results

We show our results using the above approaches in three parts. First, in Section 5.1, we define treatment status by benefits receipt and estimate differential changes in the probability of having a child for those on benefits with two or more existing children. This has the benefit of enabling us to accurately track fertility trends for those actually treated by the policy, albeit with potential bias and endogeneity. In Section 5.2, we adopt more of an intention-to-treat approach and use the fuzzier proxy of low-income occupation to estimate differential changes in the probability of having a child for those in low-income occupations. In Section 5.3, we rely exclusively on the administrative births data to estimate the impact of the policy on the probability that a child, having been born, is in the treatment group (i.e., a higher-order (3+) birth to a low-income family). Taken together, these results together give a consistent picture of the impact of the two-child limit on higher-order fertility.

5.1 Differential changes in the probability of having a child, by benefits receipt and family size

Figure 2 charts the probability of having a child by benefits receipt, family size and year from 2015 to 2019, with simple DDD estimates shown in Table 3. Since the policy was introduced in April 2017, for these annual estimates it is not altogether clear whether to include 2017 in the pre or the post period. To allow an equal 2-year window either side of the introduction of the policy, here we omit 2017 by categorising 2015 and 2016 as the pre period and 2018 and 2019 as the post period. This enables us to look at the 'settled' impacts of the policy. While there is a fall in the probability of having a higher-order child among women on benefits in the post period, it is of the same magnitude as the fall in the probability of having a lower-order child among women on benefits. The mechanical effect of the triple difference is to produce a slight positive effect (irrespective of whether 2017 is included in the post period). It is implausible to suggest that the policy increased fertility among the treated group, so we interpret this as consistent with a null hypothesis of no effect.

5.2 Differential changes in the probability of having a child, by occupation and family size

Next, we estimate the triple differences using a fuzzier but more exogenous measure of the likelihood of being treated by the two-child limit: low-income occupation. Figure 3 and Table 4 show the main results from this approach. The change in the probability of having a third or subsequent child conditional on being in a low-income occupation is not dissimilar from that shown in Figure 2. But unlike in the previous analysis, other groups do not share this reduction. Consequently, the overall triple differences estimate is negative and suggests a reduction in the probability of having a child of approximately 0.36 percentage points (when we include 2017 in the post period, there is a reduction of 0.19 percentage points). This suggests that our previous analysis, based on benefits receipt, may suffer from some bias due to compositional changes in who reports receiving benefits over time. The APS suggests that 1,551,482 women are in the treatment group (as defined here: women in families with low-income occupations with 2+ children) each year, on average. The changes in probability in Table 4 therefore imply a reduction in the number of births of approximately 17,000 a year, approximately 5600 a year of which can be attributed to the two-child limit. This equates to just under 1 percent of total annual births in England and Wales.

5.3 Differential changes in the probability of being a higher-order birth, by occupation and local deprivation

We now focus on the population of births that did take place over the period in question, using administrative births data alone to detect whether there are differential changes in the probability of a child, having been born, being a higher-order (3+) birth to a low-income family. Figure 5 charts trends in higher-order births (i.e., third or subsequent births) for our treatment and control groups (low-income occupations and higher-income occupations, respectively). As in the previous section, this is a version of a triple difference analysis. If the two-child limit had reduced higher-order fertility, we would expect the line for the treatment group (low-income occupations) to decline more rapidly than the line for the control group (higher-income occupations) after the cut-off. This does not appear to be the case. However, it is possible that the effects are not immediately visible from descriptive analysis, and/or that they are being obscured by the broader trends described above. We turn therefore to a more formal causal analysis.

Our first identification approach (as outlined in Section 4.2) is effectively a version of the chart above, and is shown in Table 5: we compare the probability of a child born into a household in the treatment group being a third or subsequent birth relative to a first or second birth, before and after the policy's birth cut-off. Column 1 shows results for our preferred measure of the treatment group: combined NS-SEC occupations with high levels of benefits receipt (and thereby low levels of income). A child being born to a low-income occupation household in the post period appears to be 0.07 percentage points more likely to be a third or subsequent birth; this estimate is both small and of the "wrong" sign, as it is not plausible for there to be a positive effect of the two-child limit on higher-order fertility. As a robustness check, Column 2 defines treatment by the mother's occupation rather than the combined occupational status of the mother and father: results are very similar.[9] Columns 3 and 4 show results for our alternative treatment variables: the

[9]We prefer the former approach, since for couple households the probability of benefit receipt is likely to be driven by the occupation of the primary earner, not solely the mother.

local proportion of the population who are income-deprived and the local proportion of households who are affected by the two-child limit. These alternative definitions suggest negative treatment effects, but they are not statistically significant, which is notable given the extremely large sample size. Overall, these results are consistent with the results in Section 5.1 and Section 5.2: any impact on fertility was small.

The significance of the variable $Post_t$ also suggests that there was also a more general trend over time. In order to unpick these factors, we repeat the above regression, but replace the $Post_t$ dummies with individual dummy variables for each half-year of birth; this allows the impact of the treatment variable to vary, not just before and after the policy change but in each half-year of birth. The results are shown graphically below in Figure 5 for our preferred treatment variable.

This presents a somewhat more complex picture. In particular, it suggests that there may have been a modest pre-treatment upward trend in the relative proportions of higher-order births among the treatment group compared to the control group. This suggests that the parallel trends assumption – that the treatment and control group would have evolved similarly in the absence of the policy – is not satisfied in this model. Nevertheless, the shape of the trend is informative. Had the previous trend continued, higher-order fertility in the treated group would have been somewhat higher; arguably, this is at least suggestive of an impact of the policy change.

However, it is worth noting that this is not, in the triple differences framework, driven by an acceleration in the decline of the proportion of higher-order births to women in low-income occupations; instead, it is driven by a deceleration in the decline of the proportion of higher-order births to women in high-income occupations, as can be seen in Figure 4 above. In other words, the change in trend occurred in the control group, not the treatment group. To attribute this to the policy implies that under the counterfactual assumption of no policy change, a similar change in trend would have occurred in the treatment group. While this is certainly possible, it is also the case that other factors could also have driven this trend break; arguably, attributing this to the policy change could simply be an artefact of the triple difference framework.

Nevertheless, we test this model formally in Table 6, by adopting an alternative specification that allows for linear time trends in the probability of being a higher-order birth and in the impact of the treatment variable, and for the latter trend to change for the treated group at the point of policy change. The coefficients of interest here are β_4, which tests for a step change in the relative level of higher-order fertility of the treatment group at the point of the policy change, after controlling for time trends; and β_7, which tests for a change in slope of the (previously upward) trend in the relative fertility of the treatment group. In other words, we are now testing for the impact of the policy change on both the level and the slope of the impact of being in the treated group on the relative probability of being a higher-order birth, after controlling for existing impacts on levels and slopes.

Our results confirm that while there was no step-change reduction in (relative) higher-order fertility due to the two-child limit, there was a reduction in the trend (or slope) after the policy cut-off. This supports the graphical evidence of Figure 4; in particular, Column 1 suggests that the policy led, for the treatment group, to a reduction in the probability of being a higher-order birth of 1 percentage point per year on average (or 3.4 percent relative to the treatment mean). After 2 years this would translate into a reduction from about 33 percent to 31 percent. This is quite similar to our preferred simple triple-difference estimate from Section 5.2, and would imply a reduction of about 6,000 births per year.

Our alternative treatment variables produce estimates that are consistent with this in sign and magnitude. Column 2 suggests that a 1 percentage point increase in the share of the local population who are income-deprived is associated with a 0.06 percentage point decline in the probability of being a higher-order birth per year due to the policy; since the average for this variable is approximately 18.4, this implies a similar size effect (of approximately 1.1 percent) to our preferred estimate. Column 3 similarly indicates that a 1 percentage point increase in the proportion of the local population affected by the two-child limit is associated with a reduction in the probability of being a higher-order birth of approximately 1.4 percentage points per year after the policy's introduction. With the average for this variable is being 1.24, this similarly applies an effect size of approximately 1.1 percent.

We test the robustness of our identification strategy in this model by performing placebo cut-off tests at dates prior to the introduction of the two-child limit (see Section 7.2 of the Appendix for full results). All three treatment variables perform well under this test, with the placebo cut-offs indicating no significant change in trend at the announcement of the policy.

Consistent with simple, standard triple difference analysis reported in Section 5.2, this analysis implies that there may have been an effect of the policy in moderating an existing pre-existing trend towards a greater share of higher-order births being to women from lower SES groups. We regard this as supporting evidence for the results found with the standard approach; that is, that the policy is likely to have reduced fertility among those affected, but the magnitude of such impacts are, at most, fairly small.

6 Discussion

While interpretation of our analyses is complicated by the pre-existing trends we observe in the data, in particular the reduction in overall fertility, taken together our conclusion is that the introduction of the two-child limit had a measurable, but relatively small impact on the number of births to women in households who were affected by the limit.

Our preferred estimate – the triple differences analysis by occupation in Section 5.2 – suggests a reduction in the probability of having a child of 0.36 percentage points (4.8 percent in relative terms). We can compute an implied elasticity of fertility with respect to income, by comparing this to the level of benefit cap (the level at which total benefits are capped in the UK, which is 20,000 GBP a year for most of the country). This implies that the loss of income (compared to the previous system) for parents choosing to have a third child is approximately 14 percent. [10] The implied elasticity is therefore 0.34. By contrast, the elasticity implied by Brewer, Ratcliffe & Smith (2012) - which found an increase in fertility among those affected of more than 15 percent in response to an income increase of about 10 percent - is closer to 1.5. Other papers - notably Milligan (2005) whose triple difference identification strategy is closest to our own - find even larger elasticities, with considerably larger fertility responses in response to much smaller changes in family income.

This result appears on the face of it to be surprising. As discussed above, the most closely related research suggests that increases in child-related benefits led to more-or-less immediate, significant, and quite large,

[10] In reality, average household income among those affected by the two-child limit is likely to be higher than this. Our estimation of an elasticity here is therefore an upper bound.

increases in fertility among affected groups, even though (in the case of the changes examined by Brewer, Ratcliffe & Smith (2012)) the increases were quite widely spread and some of the impacts were potentially ambiguous. Our results, by contrast, imply that large cuts, with (in theory) large and unambiguous impacts on the financial incentives to have children for the affected group, had at most small and gradual impacts. What explains this asymmetry?

Establishing the causal mechanisms at work here is beyond the scope of this paper. However, it is worth considering possible candidates. To the extent that those impacted did not know about the policy change, the impact of financial incentives would be attenuated. Recent qualitative research found that there is a lack of awareness of the policy; approximately half of the participants affected by the two-child limit did not know about the policy when they conceived their third or subsequent child (Patrick & Andersen forthcoming). The nature of the qualitative sample – which was restricted to larger families affected either by the two-child limit or the benefit cap – means that there is likely to have been some selection bias (since women who had been deterred from having another child by the policy will have been excluded), but nevertheless these results do suggest that imperfect information may have contributed towards the limited fertility response. This is consistent with broader research suggesting that levels of information among welfare recipients about the structure of the welfare system are relatively low, and that this may act to moderate the impact of financial incentives (Card 2020).

A lack of awareness of the policy prior to conception directly reduces the likelihood of a significant fertility response. In the absence of information, families may not find out about the policy until they notify the tax-benefit authorities of the new child, and their change of circumstances, thus removing the possibility of a fertility response. In some cases, families may find out about the policy during their pregnancy if they have contact with advisers or civil servants, as documented in the emergent qualitative evidence base (Patrick & Andersen forthcoming).

Where a lack of information means that families only find out about the policy during a pregnancy, abortion is the only available fertility response. There are likely to be low numbers of people who become aware of the policy during their pregnancy and are willing to have an abortion directly due to its existence. A survey of women who had abortions during the pandemic suggested that 57 percent who were likely to be affected by the two-child limit said it was a relevant factor in their decision (British Pregnancy Advisory Service 2020). However, this was a small survey with a highly selected sample and as such we cannot draw substantive causal conclusions from it. Abortion microdata is very difficult to access in the UK and published data do not include a breakdown by both the number of previous live births and by socio-economic status. Nevertheless, we examined published data on abortions by the number of previous live births in England and Wales (Department of Health and Social Care 2021). There is no evidence in the data of a substantial abortion response among those with two or more children; existing trends (towards more abortions, and a shift in the age distribution of abortions towards older women) do not appear to have shifted substantially at the time of the policy change, as shown in Figure 2 in the Supplementary Material.

However, it is worth noting that awareness of the policy may grow over time, especially as the number of affected families rises. This may increase the scale of the fertility response in the long run.

Other possible explanations are more speculative. It may be that, in contrast to Laroque & Salanié (2014)'s findings in France, in the UK women considering having a third or subsequent child are *less* responsive to economic incentives in fertility decisions. Religious and cultural factors may be relevant here: the two-child limit disproportionately affects orthodox Jewish and Muslim families (Work and Pensions Select Committee 2019). Several of the affected families in the qualitative study had religious beliefs which meant that they reported the two-child limit as having no impact on their conception decisions; adherence to their religious faith to them was more important than the financial incentives at play (Patrick & Andersen forthcoming). Cohen, Dehejia & Romanov (2013) found that the ultra-Orthodox Jewish population in Israel were less responsive in fertility behaviours to changes in financial incentives – namely a benefit cut affecting larger families.

Finally, it is of course well established in low-income countries with high fertility and high infant mortality that increased prosperity reduces fertility, and that decreased prosperity can increase it (Kleven & Landais 2017). The broad explanation here is that increased prosperity reduces infant mortality, so women need fewer children to be assured that one or more will reach adulthood; it also increases women's choice and agency, so that they can choose to have fewer children. The infant mortality effect is not normally relevant in developed economies. However, it is at least possible that even in the UK reducing access to economic resources has negative impacts on choice and agency, resulting in reduced access to contraception, worse mental health, and less interaction with health services; all of these could potentially increase fertility and thereby attenuate the response to financial incentives (Cesur, Gunes, Tekin & Ulker 2021, Kearney & Levine 2009).

7 Conclusion

Over the last thirty years, welfare states in the developed world have become concerned with the potential for moral hazard within the context of welfare and fertility: the assumption that expansions in benefits lead to expansions in fertility, and vice versa. In the 1990s, this motivated several US states to introduce 'family caps' under Clinton's programme of welfare reform; today it can be seen in the handful of countries that cap means-tested child benefits at the third or fourth child (Longman, Patrick, Reeves & Stewart forthcoming). However, until 2017 no country has attempted to cap child benefits at the second child (Longman, Patrick, Reeves & Stewart forthcoming). Identifying the causal impact of the UK's two-child limit on fertility therefore offers a unique opportunity to identify the causal impact of capping child benefits by family size. It also offers opportunities for credible causal identification through a triple differences design.

In this paper we show that capping child benefits at the second child leads to much smaller fertility effects than one might expect. This is surprising given the literature to date, which through the examination of benefits expansions has suggested relatively large elasticities between benefits and fertility (Brewer, Ratcliffe & Smith 2012, Laroque & Salanié 2014). We are not able to identify the reasons behind this asymmetry, but we speculate that imperfect information about the policy, the relative 'stickiness' of attitudes towards abortion (and indeed fertility preferences more generally), and the negative effect of benefit cuts on choice and agency, may have been important.

What are the policy implications of this? If capping child benefits does not have large impacts on fertility, then the implication of our findings is that the savings from the policy result almost exclusively from lower payments to poorer families, with only a marginal additional impact from reduced fertility. In having greater household needs and lower work intensity on average, larger families already faced a disproportionate risk of poverty, and child poverty in the UK among this group has increased sharply since 2013-14 (Bradshaw 2020, Stewart, Reeves & Patrick 2021). Since the two-child limit does not appear to have changed fertility behaviour or the number of births in larger families, it appears inevitable that it will increase child poverty further among larger families. This has significant implications for inequalities in children's outcomes and development (Cooper & Stewart 2017).

Finally, in implying an asymmetry between the effects of benefits expansions and benefits reductions, our results have wider policy implications in showing that it cannot simply be assumed that doing the opposite of a policy will lead to equivalent results in the opposite direction. In our context, the government understood and framed the two-child limit as a policy that would have the opposite effects of previous benefits expansions: it would reduce fertility, possibly by a substantial amount (Treasury & DWP 2015). Yet our results suggest a relatively small impact on fertility. This underscores the need for robust causal evidence specific to the policy in question during the policymaking process.

References

Becker, G. (1960), *An Economic Analysis of Fertility*, Demographic and Economic Change in Developed Countries, a Conference of the Universities, National Bureau Committee for Economic Research, Princeton University Press, Princeton. Last accessed: 03/02/22.

Bradshaw, J. (2020), 'The two-child limit: impact on abortion', https://bit.ly/3GxNwb9. Last accessed: 03/02/22.

Brewer, M., Ratcliffe, A. & Smith, S. (2012), 'Does welfare reform affect fertility? Evidence from the UK', *Journal of Population Economics* **25**(1), 245–266.

British Pregnancy Advisory Service (2020), 'Forced into a corner: the two-child limit and pregnancy decision-making during the pandemic', https://bit.ly/35FXzhz. Last accessed: 03/02/22.

Camasso, M. J. & Jagannathan, R. (2016), 'The future of the family cap: fertility effects 18 years post-implementation', *Social Service Review* **90**(2), 264–304.

Card, D. (2020), 'Reforming the financial incentives of the welfare system', https://davidcard.berkeley.edu/papers/finan-incen-welfare.pdf. Last accessed: 03/02/22.

Cesur, R., Gunes, P. M., Tekin, E. & Ulker, A. (2021), 'Socialized healthcare and women's fertility decisions', *Journal of Human Resources* p. 419.

Cohen, A., Dehejia, R. & Romanov, D. (2013), 'Financial incentives and fertility', *The Review of Economics and Statistics* **95**(1), 1–20.

Cooper, K. & Stewart, K. (2017), 'Does money affect children's outcomes? an update', *Centre for Analysis of Social Exclusion* .

Department of Health and Social Care (2021), 'Abortion statistics 2020: freedom of information request and other data releases', https://bit.ly/3guQDWH. Last accessed: 03/02/22.

Department of Work and Pensions and Her Majesty's Revenue and Customs (2021), 'Universal credit and child tax credit claimants: statistics related to the policy to provide support for a maximum of 2 children, april 2021 and previous versions', https://bit.ly/348aw3v. Last accessed: 03/02/22.

Dyer, W. T. & Fairlie, R. W. (2004), 'Do Family Caps Reduce Out-of-Wedlock Births? Evidence from Arkansas, Georgia, Indiana, New Jersey and Virginia', *Population Research and Policy Review* **23**(5/6), 441–473.

Gaffney, D. (2015), 'Retrenchment, reform, continuity: Welfare under the coalition', *National Institute Economic Review* **231**(231), R44–R53.

Gaffney, D. & Portes, J. (2013), 'Conservative claims about benefits are not just spin, they're making it up', https://bit.ly/3ox0Rds. Last accessed: 03/02/22.

González, L. & Trommlerová, S. K. (2021), 'Cash transfers and fertility: How the introduction and cancellation of a child benefit affected births and abortions', *Journal of Human Resources* p. 220.

Grice, A. (2015), ‘Budget 2015: Tax credits and housing benefit to be cut for families with more than two children’, https://bit.ly/3ryWHUB. Last accessed: 03/02/22.

Jensen, T. & Tyler, I. (2015), ‘”benefits broods”: The cultural and political crafting of anti-welfare commonsense’, *Critical Social Policy* **35**(4), 470–491.

Joyce, T., Kaestner, R., Korenman, S. & Henshaw, S. (2004), ‘Family cap provisions and changes in births and abortions’, *Population Research and Policy Review* **23**(5/6), 475–511.

Kearney, M. S. (2004), ‘Is there an effect of incremental welfare benefits on fertility behavior? a look at the family cap’, *Journal of Human Resources* **39**(2), 295–325.

Kearney, M. S. & Levine, P. B. (2009), ‘Subsidized contraception, fertility, and sexual behavior’, *The Review of Economics and Statistics* **91**(1), 137–151.

Kleven, H. & Landais, C. (2017), ‘Gender inequality and economic development: Fertility, education and norms’, *Economica* **84**(334), 180–209.

Laroque, G. & Salanié, B. (2014), ‘Identifying the response of fertility to financial incentives’, *Journal of Applied Econometrics* **29**(2), 314–332.

Longman, G., Patrick, R., Reeves, A. & Stewart, K. (forthcoming), ‘The benefit cap and two-child limit in comparative perspective’.

Marsh, S. (2017), ‘Two-child limit receiving benefits: are you going to be affected?’, https://bit.ly/3Jauycj. Last accessed: 03/02/22.

Milligan, K. (2005), ‘Subsidizing the stork: New evidence on tax incentives and fertility’, *The Review of Economics and Statistics* **87**(3), 539–555.

Ministry of Housing, Communities and Local Government (2020), ‘Indices of deprivation 2019: income and employment domains combined for england and wales - guidance note’, https://bit.ly/3Gmk5bX. Last accessed: 03/02/22.

OECD (2021), ‘Purchasing power parities’, https://bit.ly/3JbkH5Z. Last accessed: 03/02/22.

Office for National Statistics (2016), ‘Birth registrations, england and wales microdata metadata’, https://bit.ly/3rAkYd1. Last accessed: 03/02/22.

Office for National Statistics (2021), ‘Estimated number of households by selected household types, local authorities in England and Wales, counties and regions of England, Scottish council areas, and Great Britain constituent countries, 2004 to 2019’, https://bit.ly/330hkj6. Last accessed: 03/02/22.

Patrick, R. & Andersen, K. (forthcoming), ‘The two-child limit ‘choices’ over family size: When policy presentation collides with lived experiences’.

Reader, M. & Andersen, K. (2022), Size matters: the experiences of larger families on a low income during covid-19, *in* K. Garthwaite, R. Patrick, A. Tarrant & R. Warnock, eds, ‘COVID-19 Collaborations: Researching poverty and low-income family life during the pandemic’, Policy Press, Bristol.

Reed, H. & Portes, J. (2015), 'Cumulative impact assessment of tax and welfare reforms', *Equality and Human Rights Commission* . Last accessed: 03/02/22.

Riphahn, R. T. & Wiynck, F. (2017), 'Fertility effects of child benefits', *Journal of Population Economics* **30**(4), 1135–1184.

Slack, J. (2015), 'Now the taxpayer won't fund big families: Tax credits to be limited to two children to cut bill that's soared to £30billion', https://bit.ly/3J9OxIg. Last accessed: 03/02/22.

Stewart, K., Reeves, A. & Patrick, R. (2021), 'A time of need: Exploring the changing poverty risk facing larger families in the uk', *Centre for Analysis of Social Exclusion* .

Treasury, H. & DWP (2015), 'Welfare reform and work bill: Impact assessment of tax credits and universal credit, changes to child element and family element', https://bit.ly/3HD99bg. Last accessed: 03/02/22.

Treasury, H. M. (2015), 'Summer budget 2015, hc 264, paras 1.141–1.150'.

UK Government (2021), 'Pm's speech on welfare reform bill', https://bit.ly/3gm9Pps. Last accessed: 03/02/22.

Whitehouse, H. (2015), 'Budget 2015: Child tax credits to be limited to two children after 2017', https://bit.ly/3rBvMYm. Last accessed: 03/02/22.

Work and Pensions Select Committee (2019), 'The two-child limit: Third report of session 2019', https://bit.ly/3uwvTWX. Last accessed: 03/02/22.

Table 1: Summary statistics for administrative births microdata in England and Wales, 2015-2019

	N	Mean	SD
Multiple birth	3,380,560	0.030	0.171
Stillbirth	3,380,560	0.004	0.064
Previous live births	3,380,560	0.967	1.158
Maternal age (years)	3,380,510	30.530	5.651
Paternal age (years)	3,200,330	33.396	6.720
Income deprivation score	3,380,560	14.956	9.983
Local incidence of the two-child limit	3,345,233	1.240	0.583
Low-income occupation	329,594	0.491	0.500
High-income occupation	329,594	0.509	0.500
Single parent	3,380,560	0.157	0.364

Note: Income deprivation score represents the proportion of the local population who receive means-tested benefits. NS-SEC occupation is combined occupation, where the highest-ranked occupation of the mother and father is selected. Single mothers either register the birth on their own or report that the parents live at different addresses. Local incidence of the two-child limit is the percentage of the population in the child's local authority who are affected by the two-child limit.

Table 2: Annual population estimates from Annual Population Survey sample, women aged 16-45 in England and Wales, 2015-2019

Total population of interest (women aged 16-45)	11,237,692
No benefits, 0-1 child	6,989,810
No benefits, 2+ children	2,110,500
Benefits, 0-1 child	1,002,960
Benefits, 2+ children	1,134,422
High-income occupation, 0-1 child	4,534,212
High-income occupation, 2+ children	1,693,440
Low-income occupation, 0-1 child	3,458,558
Low-income occupation, 2+ children	1,551,482

Note: Figures are weighted averages across 2015, 2016, 2017, 2018 and 2019. Unweighted N=265,070. Person-household weightings used to correct for non-response. NS-SEC refers exclusively to the highest occupation status of the family. Benefits receipt includes Universal Credit, tax credits, housing benefit, Jobseeker's Allowance (JSA), and income support. Number of children refers to dependent children under 16 in the family.

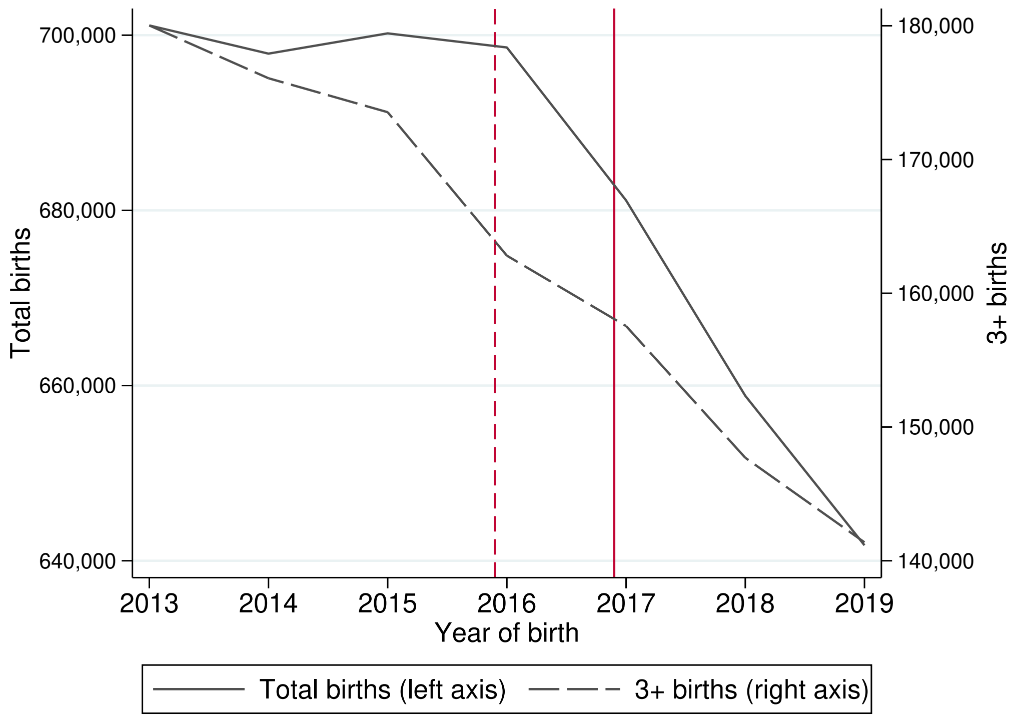


Figure 1: Number of total and 3+ births in England and Wales by year, 2013-2019

Note: Data collapsed to annual data points here, as the number of births is volatile by month and quarter of birth. Red solid line indicates the introduction of the two-child limit in April 2017 (here it is displayed just before 2017, as the data is annual and the policy affected the majority of 2017). Red dashed line indicates the announcement of the two-child limit in July 2015 (for the same reasons, here it is displayed just before 2016).

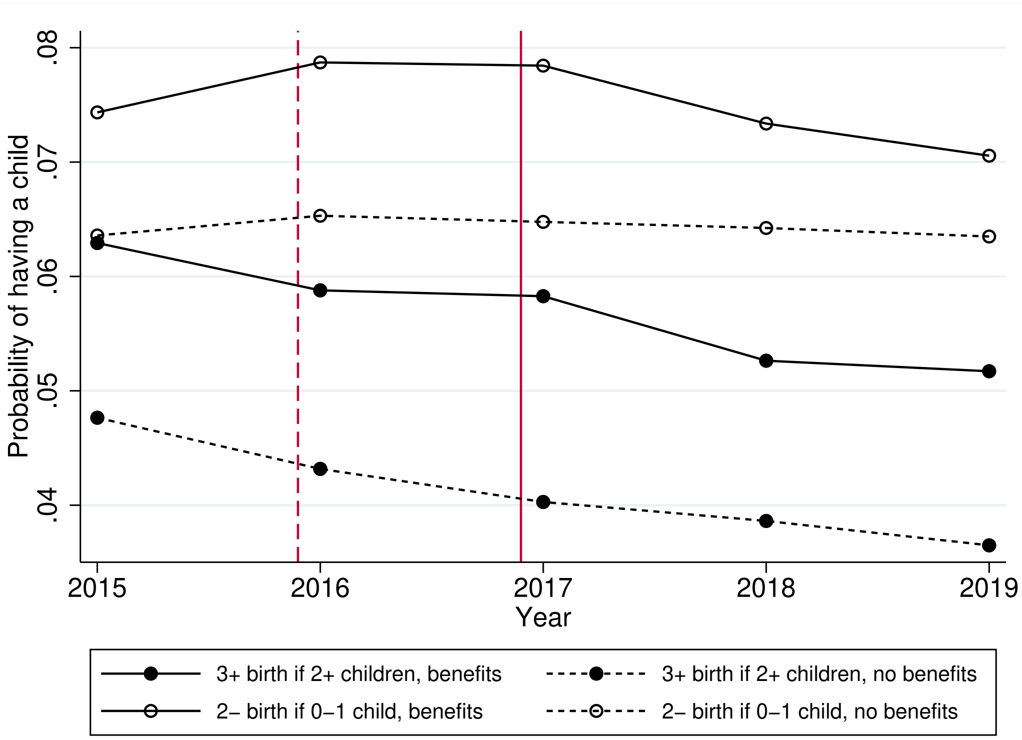


Figure 2: Probability of having a child by benefits receipt, family size and year, 2015-2019

Note: Data from Annual Population Survey and administrative births microdata. Sample is adult female respondents aged 16-45 (weighted N= 11 million a year on average; unweighted N=53,014 a year on average).

Table 3: Triple difference estimates by benefits receipt and family size

	2+ children		0-1 children	
Year	Benefits	No benefits	Benefits	No benefits
2015	0.0629	0.0476	0.0743	0.0636
2016	0.0588	0.0432	0.0787	0.0653
2017	0.0583	0.0403	0.0784	0.0648
2018	0.0526	0.0386	0.0734	0.0642
2019	0.0517	0.0365	0.0706	0.0635
Pre (2015-16)	0.0609	0.0454	0.0765	0.0644
Post (2018-19)	0.0522	0.0376	0.072	0.0639
First difference	-0.0087	-0.0079	-0.0046	-0.0006
Second difference	-0.0008		-0.004	
Third difference	0.0032			

Note: Figures show the probability of having a child for each group. Data on the number of lower- and higher-order births by NS-SEC category and year from birth records; data on benefits probabilities and the number of women aged 16-45 by NS-SEC category and family size from the Annual Population Survey. The first difference subtracts the post from the pre averages. The second difference subtracts the first differences of those on benefits from those not on benefits. The third difference subtracts the second differences of higher-order births to those with 2+ children from lower-order births to those with 0-1 child.

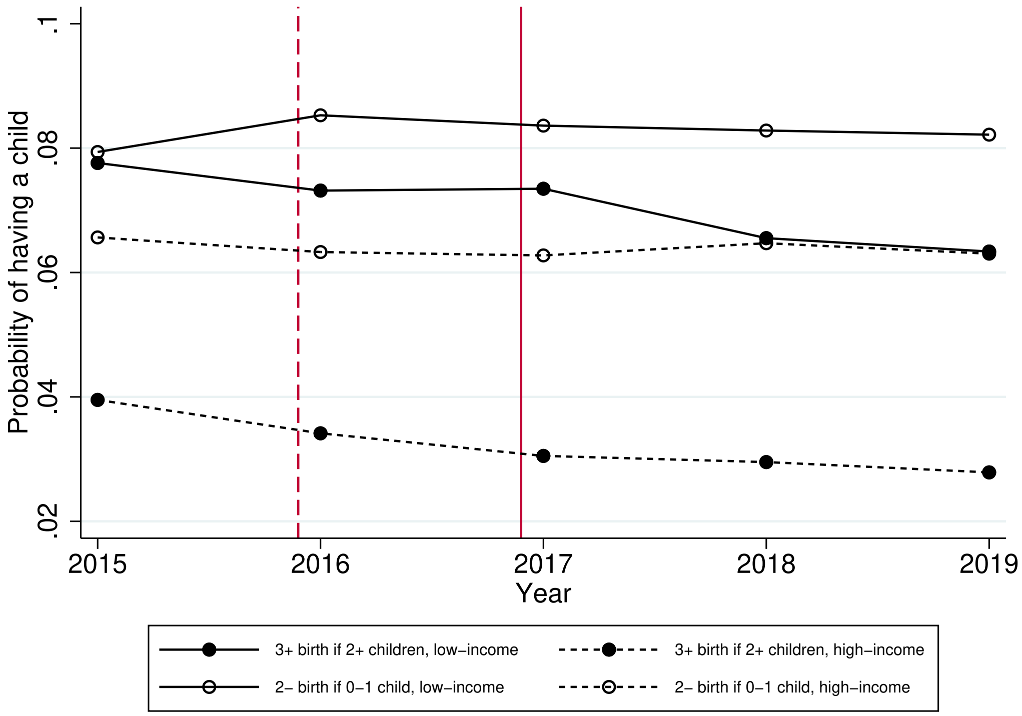


Figure 3: Probability of having a child by NS-SEC occupation, family size and year, 2015-2019

Note: Data from Annual Population Survey and administrative births microdata. Sample is adult female respondents aged 16-45 (weighted N=11 million a year on average; unweighted N=53,014 a year on average).

Table 4: Triple difference estimates by family occupation and family size

	2+ children		0-1 children	
Year	Low-income	High-income	Low-income	High-income
2015	0.0776	0.0395	0.0794	0.0656
2016	0.0732	0.0342	0.0853	0.0633
2017	0.0735	0.0305	0.0836	0.0628
2018	0.0655	0.0295	0.0828	0.0647
2019	0.0634	0.0279	0.0822	0.063
Pre (2015-16)	0.0754	0.0368	0.0823	0.0645
Post (2018-19)	0.0645	0.0287	0.0825	0.0639
First difference	-0.0109	-0.0081	0.0002	-0.0006
Second difference	-0.0028		0.0008	
Third difference	-0.0036			

Note: Figures show the probability of having a child for each group. Data on the number of lower- and higher-order births by NS-SEC category and year from birth records; data on the number of women aged 16-45 by NS-SEC category and family size from the Annual Population Survey. The first difference subtracts the post from the pre averages. The second difference subtracts the first differences of those on in low-income occupations from those in high-income occupations. The third difference subtracts the second differences of higher-order births to those with 2+ children from lower-order births to those with 0-1 child.

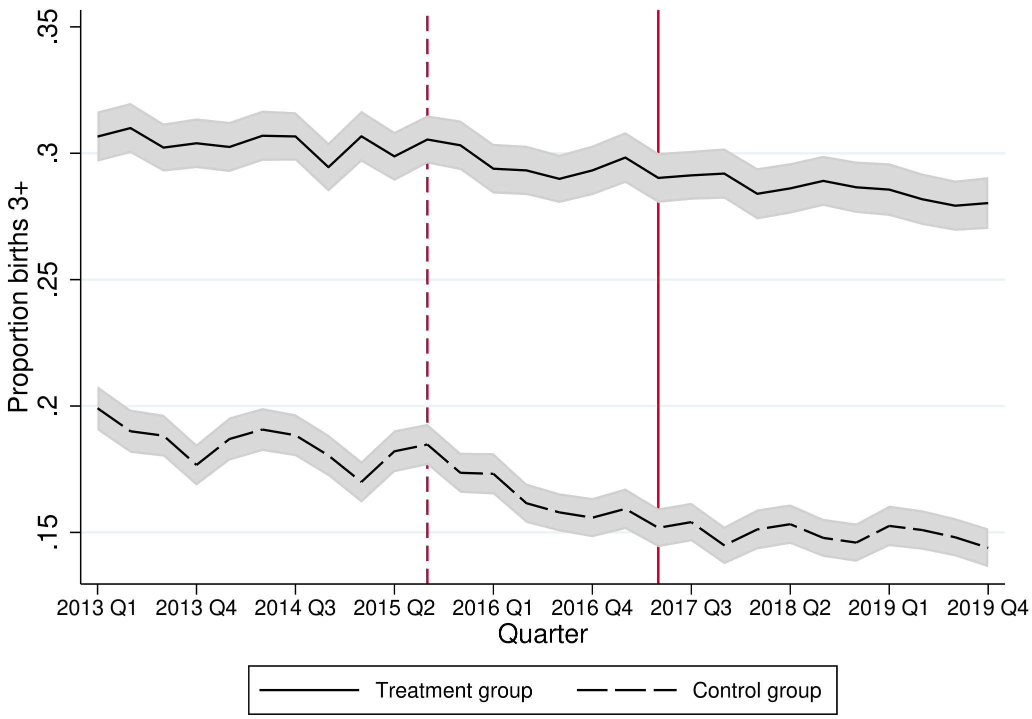


Figure 4: Proportion of total births who are higher order (third or subsequent births), by treatment and control group

Notes: Data collapsed to quarter of birth. A birth is in the treatment group if the parents are in low-income occupations (NS-SEC 3-7); it is in the control group if the parents are in higher-income occupations (NS-SEC 1-2). Grey shaded areas are 95 percent confidence intervals. Red solid line indicates the date of birth cut-off for the introduction of the two-child limit (6 April 2017). Red dashed line indicates the date of the announcement of the two-child limit (5 July 2015).

Table 5: Effects of the two-child limit on the probability of being a higher-order birth: triple differences results

	(1)	(2)	(3)	(4)
Post	-0.019***	-0.016***	-0.014***	-0.015***
	(0.002)	(0.002)	(0.001)	(0.003)
Treat	0.130***	0.062***	0.006***	0.064***
	(0.002)	(0.002)	(0.000)	(0.004)
Post*Treat	0.007**	0.011***	-0.000	-0.001
	(0.003)	(0.003)	(0.000)	(0.002)
Constant	0.169***	0.131***	0.142***	0.161***
	(0.001)	(0.001)	(0.001)	(0.004)
N	329,594	245,593	3,380,560	3,345,233

Note: Data from administrative birth registrations microdata for England and Wales, 2012-2019. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Dependent variable: Dummy variable equal to one if a birth is the third or subsequent live birth to the mother, and zero otherwise. Linear probability model. 'Post' is a dummy variable equal to one if the child is born on or after 6 April 2017 and zero otherwise. 'Treat' refers to the relevant measure of low-income, as reflected in the column titles: column 1 defines treatment as having parents in low-income occupations (NS-SEC 3-7); column 2 defines treatment as the mother being in a low-income occupation (NS-SEC 3-7); column 3 defines treatment as the percentage of the local population who are income-deprived (i.e., a continuous Index of Income Deprivation score at Lower Super Output Area (LSOA)); column 4 defines treatment as the percentage of the local population who are affected by the two-child limit (i.e., the percentage of households in the mother's local authority of residence who were affected by the two-child limit in 2021, as indicated by government published statistics (Department of Work and Pensions and Her Majesty's Revenue and Customs 2021)). Standard errors clustered at LSOA for column 3; at LA level for column 4. Robust standard errors are utilised for column 1 and 2. The number of total observations for column 1 is lower because occupation is only coded for approximately a 10 percent random sample of the births microdata, whereas geography is coded for all observations.

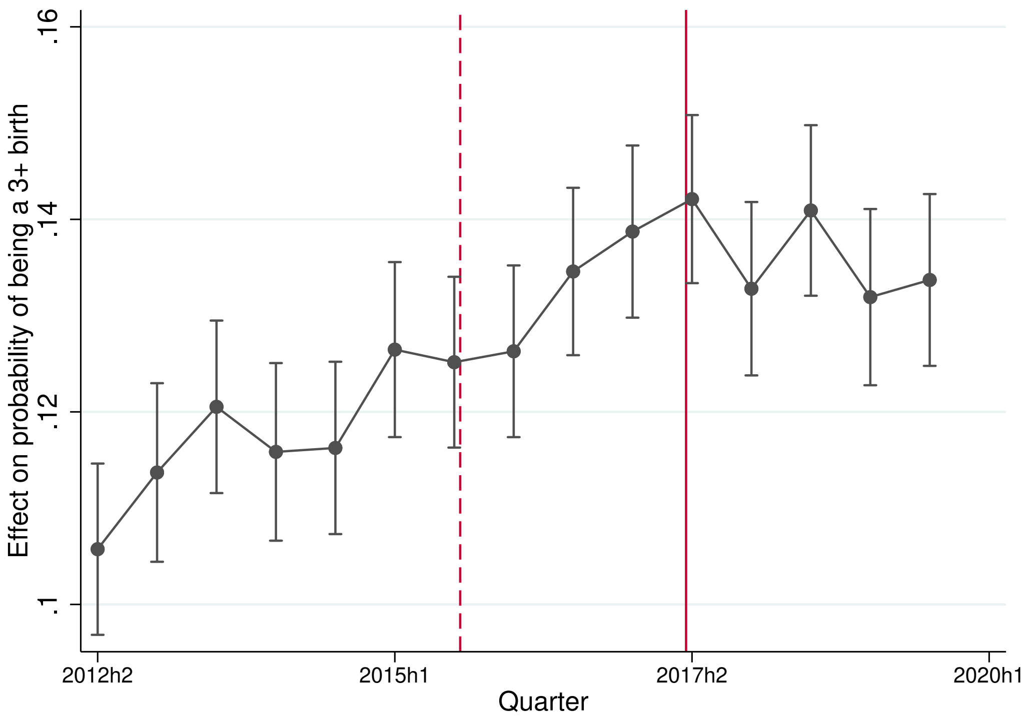


Figure 5: Effect of parents being in low-income occupations on the probability of a baby being a higher-order birth (i.e., third or subsequent birth), by half-year of birth

Notes: Data from administrative birth registrations microdata for England and Wales, 2012-2019. Markers indicate the effect of being in the treatment group on the probability of the child being a third or subsequent birth for each half-year of birth. Grey bars are 95 percent confidence intervals. Red solid line indicates the date of birth cut-off for the introduction of the two-child limit (6 April 2017). Red dashed line indicates the date of the announcement of the two-child limit (5 July 2015).

Table 6: Triple differences estimates using administrative births data, accounting for time trends

	(1)	(2)	(3)	(4)
Post	-0.003	-0.003	0.008***	0.011***
	(0.004)	(0.004)	(0.002)	(0.003)
DOB	-0.013***	-0.012***	-0.017***	-0.022***
	(0.002)	(0.002)	(0.001)	(0.002)
Post*DOB	0.011***	0.011***	0.014***	0.021***
	(0.003)	(0.003)	(0.001)	(0.003)
Treat	0.138***	0.064***	0.007***	0.074***
	(0.004)	(0.004)	(0.000)	(0.004)
Post*Treat	0.002	0.006	-0.000***	-0.006**
	(0.006)	(0.006)	(0.000)	(0.003)
Treat*DOB	0.007**	0.002	0.000***	0.009***
	(0.003)	(0.003)	(0.000)	(0.002)
Post*Treat*DOB	-0.010**	-0.001	-0.001***	-0.013***
	(0.004)	(0.004)	(0.000)	(0.002)
Constant	0.154***	0.118***	0.123***	0.137***
	(0.003)	(0.003)	(0.001)	(0.004)
N	329,594	245,593	3,380,560	3,345,233

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10. Dependent variable: Dummy variable equal to one if a birth is the third or subsequent live birth to the mother, and zero otherwise. Linear probability model. 'Post' is a dummy variable equal to one if the child is born on or after 6 April 2017 and zero otherwise. DOB_t is the child's date of birth, centred at the cut-off and expressed in years[11] 'Treat' refers to the relevant measure of low-income, as reflected in the column titles: column 1 defines treatment as having parents in low-income occupations (NS-SEC 3-7); column 2 defines treatment as the mother being in a low-income occupation (NS-SEC 3-7); column 3 defines treatment as the percentage of the local population who are income-deprived (i.e., a continuous Index of Income Deprivation score at Lower Super Output Area (LSOA)); column 4 defines treatment as the percentage of the local population who are affected by the two-child limit (i.e., the percentage of households in the mother's local authority of residence who were affected by the two-child limit in 2021, as indicated by government published statistics (Department of Work and Pensions and Her Majesty's Revenue and Customs 2021)). Standard errors clustered at LSOA for column 3; at LA level for column 4. Robust standard errors are utilised for column 1 and 2. The number of total observations for column 1 is lower because occupation is only coded for approximately a 10 percent random sample of the births microdata, whereas geography is coded for all observations.

Appendix A. Robustness Checks

7.1 Anticipation or lagged effects

While there do not appear to be large fertility effects at the introduction of the two-child limit, we consider the possibility that there may have been anticipation effects after the announcement of the policy in 2015, or lagged effects after the policy's introduction.

The policy was announced in the July 2015 budget for implementation in 2017. Some media coverage, particularly in the broadsheet press, stated clearly that the policy would be introduced in 2017 and that 'families who have a third child after April 2017 could be caught' (Grice 2015, Slack 2015). However, other media sources made no such clarification, meaning that some families may not have known about the birth cut-off (Whitehouse 2015). On the (albeit unlikely) assumption that parents had perfect information about the policy announcement, this would give families almost two years to respond. Parents who were considering having a third child might decide to do so quickly in advance of the birth cut-off, so as not to ensure they receive benefits for that child. Alternatively, parents may interpret the policy as a wider signifier of a lack of support for larger families and decide not to have a third child. The direction of these anticipation effects is therefore not clear in theory.

It is also possible that fertility effects may have been lagged: parents may only have become aware of the policy once it was rolled out and once greater media attention was drawn to the policy (Marsh 2017).

To test for anticipation or lagged treatment effects, we examine our multiple period estimates to see whether there are notable trends before or after the policy's formal introduction. As Figure 4 shows, there is no evidence of anticipation effects after the announcement of the policy and before the implementation of the policy (i.e., between the dashed and solid vertical lines). Neither is there evidence for a lagged fertility response so far.

7.2 Placebo cut-off tests

We conduct placebo cut-off tests to test the validity of our identification strategy. We pretend that the two-child limit was introduced on placebo dates prior to its introduction, and test whether our methodology indicates any spurious results. We test for placebo effects on three dates: 6 April 2015, two years prior the introduction of the policy and prior to the announcement of the policy; 8 July 2015, the exact date on which the two-child limit was announced, when effects should be not be possible; and 6 April 2016, one year prior to the policy's introduction but after it had been announced. In all cases, we test for a two-year window, with one year either side of the cut-off. This ensures that none of the placebo samples are contaminated by the *Post* period. If our identification strategy is sound, these models should not show any significant effects, though it is possible that the 6 April 2016 date could reflect anticipation effects.

As Table 7, Table 8 and Table 9 show, there are no statistically significant step changes or changes in trend at any of these placebo dates for any of our treatment variables. This suggests that the change in trend identified in Section 5.2 is likely to be linked to the introduction of the two-child limit. It also demonstrates that there were no anticipation effects of the two-child limit, a crucial assumption for our Triple Differences identification strategy.

Table 7: Placebo cut-off test: treatment defined by low-income occupations

	(1) 08jul2015	(2) 06apr2015	(3) 06apr2016
Post	-0.023	0.034	0.004
	(0.019)	(0.021)	(0.012)
DOB	-0.008	-0.025***	-0.016**
	(0.007)	(0.007)	(0.007)
Post*DOB	-0.019*	0.009	0.013
	(0.010)	(0.010)	(0.010)
Treat	0.136***	0.171***	0.126***
	(0.025)	(0.029)	(0.017)
Post*Treat	0.002	-0.046	0.016
	(0.029)	(0.034)	(0.018)
Treat*DOB	0.007	0.020*	0.003
	(0.011)	(0.011)	(0.011)
Post*Treat*DOB	0.003	-0.017	0.009
	(0.016)	(0.016)	(0.016)
Constant	0.162***	0.119***	0.154***
	(0.017)	(0.018)	(0.011)
N	136,218	135,772	136,038

Note: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Dependent variable: Dummy variable equal to one if a birth is the third or subsequent live birth to the mother, and zero otherwise. Linear probability model. 'Post' is a dummy variable equal to one if the child is born on or after 6 April 2017 and zero otherwise. 'Treat' is a dummy variable equal to one if a birth has parents in low-income occupations (NS-SEC 3-7). Robust standard errors are utilised. The number of total observations in this model is lower than Table 8 and Table 9 because occupation is only coded for approximately a 10 percent random sample of the births microdata, whereas geography is coded for all observations.

Table 8: Placebo cut-off test: treatment defined by local deprivation

	(1) 08jul2015	(2) 06apr2015	(3) 06apr2016
Post	-0.0204**	0.0290***	0.0064
	(0.0082)	(0.0094)	(0.0051)
DOB	-0.0101***	-0.0277***	-0.0212***
	(0.0032)	(0.0032)	(0.0031)
Post*DOB	-0.0158***	0.0065	0.0123***
	(0.0044)	(0.0045)	(0.0044)
Treat	0.0073***	0.0071***	0.0069***
	(0.0004)	(0.0005)	(0.0003)
Post*Treat	-0.0004	-0.0002	0.0001
	(0.0005)	(0.0006)	(0.0003)
Treat*DOB	0.0006***	0.0005**	0.0004**
	(0.0002)	(0.0002)	(0.0002)
Post*Treat*DOB	-0.0003	-0.0001	0.0001
	(0.0003)	(0.0003)	(0.0003)
Constant	0.1342***	0.0906***	0.1196***
	(0.0072)	(0.0082)	(0.0048)
N	1,402,227	1,398,763	1,395,731

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10. Dependent variable: Dummy variable equal to one if a birth is the third or subsequent live birth to the mother, and zero otherwise. Linear probability model. 'Post' is a dummy variable equal to one if the child is born on or after 6 April 2017 and zero otherwise. 'Treat' is the percentage of the local population who are income-deprived (i.e., a continuous Index of Income Deprivation score at Lower Super Output Area (LSOA)). Standard errors clustered at LSOA level.

Table 9: Placebo cut-off test: treatment defined by local two-child limit incidence

	(1) 08jul2015	(2) 06apr2015	(3) 06apr2016
Post	-0.0306**	0.0136	0.0016
	(0.0149)	(0.0159)	(0.0081)
DOB	-0.0103*	-0.0224***	-0.0220***
	(0.0059)	(0.0060)	(0.0052)
Post*DOB	-0.0212**	0.0003	0.0151**
	(0.0085)	(0.0076)	(0.0069)
Treat	0.0704***	0.0601***	0.0667***
	(0.0112)	(0.0107)	(0.0051)
Post*Treat	0.0013	0.0066	0.0061
	(0.0105)	(0.0111)	(0.0048)
Treat*DOB	0.0074*	0.0027	0.0049
	(0.0043)	(0.0043)	(0.0032)
Post*Treat*DOB	0.0003	0.0023	0.0003
	(0.0060)	(0.0052)	(0.0044)
Constant	0.1573***	0.1267***	0.1402***
	(0.0140)	(0.0144)	(0.0083)
N	1,387,560	1,384,157	1,381,223

Note: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Dependent variable: Dummy variable equal to one if a birth is the third or subsequent live birth to the mother, and zero otherwise. Linear probability model. 'Post' is a dummy variable equal to one if the child is born on or after 6 April 2017 and zero otherwise. 'Treat' is the percentage of the local population who are affected by the two-child limit (i.e., the percentage of households in the mother's local authority of residence who were affected by the two-child limit in 2021, as indicated by government published statistics (Department of Work and Pensions and Her Majesty's Revenue and Customs 2021)). Standard errors clustered at LA level.